



# The Last Journey. II. SMACC—Subhalo Mass-loss Analysis Using Core Catalogs

Imran Sultan<sup>1,2,5</sup> , Nicholas Frontiere<sup>2,3</sup>, Salman Habib<sup>2,3</sup>, Katrin Heitmann<sup>2</sup> , Eve Kovacs<sup>2</sup> , Patricia Larsen<sup>2,3</sup>, and Esteban Rangel<sup>4</sup>

<sup>1</sup> Department of Physics & Astronomy and Center for Interdisciplinary Exploration and Research in Astrophysics, Northwestern University, 1800 Sherman Avenue, Evanston, IL 60201, USA

<sup>2</sup> High Energy Physics Division, Argonne National Laboratory, Lemont, IL 60439, USA

<sup>3</sup> Computational Science Division, Argonne National Laboratory, Lemont, IL 60439, USA

<sup>4</sup> Argonne Leadership Computing Facility, Argonne National Laboratory, Lemont, IL 60439, USA

Received 2020 December 16; revised 2021 March 16; accepted 2021 April 4; published 2021 June 1

## Abstract

This paper introduces Subhalo Mass-loss Analysis using Core Catalogs (SMACC). SMACC adds a mass model to substructure merger trees based on halo “core tracking.” Our approach avoids the need for running expensive subhalo finding algorithms and instead uses subhalo mass-loss modeling to assign masses to halo cores. We present details of the SMACC methodology and demonstrate its excellent performance in describing halo substructure and its evolution. Validation of the approach is carried out using cosmological simulations at significantly different resolutions. We apply SMACC to the 1.24-trillion-particle Last Journey simulation and construct core catalogs with the additional mass information. These catalogs can be readily used as input to semianalytic models or subhalo abundance matching approaches to determine approximate galaxy distributions, as well as for in-depth studies of small-scale structure evolution.

*Unified Astronomy Thesaurus concepts:* [N-body simulations \(1083\)](#); [Cosmology \(343\)](#); [Large-scale structure of the universe \(902\)](#)

## 1. Introduction

Cosmological simulations have provided, and continue to provide, otherwise wholly inaccessible insights into the highly nonlinear process of large-scale structure formation in the universe (for reviews, see, e.g., Bertschinger 1998; Dolag et al. 2008; Frenk & White 2012). The volume and resolution achieved by these simulations have improved tremendously over the past few decades following the trajectory for computational power set by Moore’s law. This has allowed for the emergence of an ever more detailed picture of the complex structures formed by the action of gravity in an expanding universe.

In the fiducial  $\Lambda$ CDM cosmology, structure formation is hierarchical, with small halos—localized clumps of matter—forming first and then merging into larger halos. They then, in turn, merge into even larger ones, a process captured by the notion of merger trees (Lacey & Cole 1993). Such a dynamical evolution naturally introduces the concept of halo substructure as smaller halos fall into more massive objects and live out a (potentially perilous) existence as subhalos, subject to tidal disruption and mergers. Further substructure exists within subhalos as well, complicating the analysis of high-resolution simulations.

If one is interested only in the density field associated with halos, there is no particular reason to be concerned about identifying which region is a piece of the main halo and which corresponds to substructure, as long as the density field is sufficiently resolved. More information, however, is required for further detailed analyses: investigations of cosmological probes based on large-scale structure require modeling galaxy locations, as well as other properties of galaxies. Therefore, in the context of a gravity-only  $N$ -body simulation, it becomes necessary to model galaxy formation and evolution by using detailed spatiotemporal knowledge of the properties of halos,

halo substructure, and the halo environment. Such an approach, referred to as the “galaxy–halo connection,” can be studied via various combinations of  $N$ -body simulation results, empirical modeling (e.g., halo occupation distribution (HOD), subhalo abundance matching (SHAM)), and semianalytic modeling (SAM) techniques, including modeling of gas dynamics and astrophysical feedback effects (for reviews, see Mo et al. 2011; Somerville & Davé 2015; Wechsler & Tinker 2018).

In this picture, the association of galaxies with subhalos is nontrivial because of the complexity of the ever-evolving picture of “structures within structures.” As one example, observations of group- and cluster-scale halos have shown that the radial distribution of galaxies does not follow that of simulated subhalos (Lin et al. 2004; Budzynski et al. 2012). For small-scale clustering measurements, related challenges have been discussed in Campbell et al. (2018). The one-to-one association of galaxies with their host subhalos can be lost owing to physical mechanisms such as tidal stripping. Numerical effects in simulations, due to insufficient mass resolution and problems with subhalo finding, can lead to artificial subhalo loss (especially near halo centers). Galaxies that are no longer associated with their dark matter subhalo are termed “orphan” galaxies; dealing properly with orphans is an important aspect of semianalytic galaxy modeling.

The short discussion above points to the importance of robustly tracking substructure in cosmological simulations. Identification of substructure is a challenging computational task (Muldrew et al. 2011), however, and is made more difficult by the lack of a rigorous definition of halos and their substructure and by the lack of direct observational analogs. Indeed, there are many approaches to subhalo identification and tracking in simulations, all of which have their own features and strengths and weaknesses (for a detailed comparison project, see Onions et al. 2012). In addition, the requirements on the force and mass resolution for the simulations to ensure

<sup>5</sup> Corresponding author.

that subhalos are captured reliably pose a major challenge (see, e.g., Guo & White 2014; van den Bosch 2017; van den Bosch & Ogiya 2018 for recent discussions).

In order to address some important aspects of substructure identification and the related problem of galaxy–substructure association, we have recently introduced the concept of “core tracking” in  $N$ -body simulations (Rangel et al. 2016; Heitmann et al. 2021; D. Korytov et al. 2021 in preparation). The intellectual ancestry of this idea goes back to Kaiser (1984) and White et al. (1987), focusing on following substructure peaks as potential galaxy locations. In our present-day context, a core is defined by identifying a set of particles closest to the center of a halo, the process beginning when a halo reaches some initial mass threshold. Once found, all core particles are tracked for the remainder of the simulation. By combining the core particle information with detailed halo merger trees, one can build core catalogs that follow the cores from their birth to the final step in the simulation. Two advantages of core tracking that address the aforementioned difficulties of subhalo finding are that (1) cores are never lost because they are tracked forever and (2) core tracking requires only list construction and comparison, making the approach computationally inexpensive compared to subhalo-based methods.

The core-tracking approach has already been successfully applied in D. Korytov et al. (2021, in preparation), demonstrating the capability to model dynamical processes such as galaxy mergers and disruption, thereby circumventing the problem of missing orphan galaxies inherent with subhalo finding. However, core tracking is by no means a complete description of substructure: once a halo has merged into another halo and become an internal structure, its identity becomes uncertain and its properties are not easily measurable—tracking the original halo core is, at first sight, not obviously of direct help in reconstructing, for example, the evolution of the associated subhalo mass. Fortunately, previous work has shown that the evolution of subhalo masses in simulations appears, on average, to follow simple analytic forms that characterize the mass loss as the subhalos orbit around a host object (van den Bosch et al. 2005; Giocoli et al. 2008; Jiang & van den Bosch 2016; van den Bosch & Jiang 2016; Han et al. 2018; Hiroshima et al. 2018). While the reality is complicated by other dynamics, van den Bosch et al. (2005) have been able to reproduce the subhalo mass function (SHMF) from initial substructure masses originally developed in the context of extended Press–Schechter (EPS) trees. Building on these studies, we can determine a simple mass-loss model for predicting the evolution of the subhalo mass associated with merged cores.

In this paper, we introduce SMACC—Subhalo Mass-loss Analysis using Core Catalogs—which allows us to build core catalogs that not only contain the full merger history for each core but also model the evolution of its associated mass beyond the time it has merged with a more massive object. The resulting “core masses” can be used as inputs to SAMs to create synthetic galaxy catalogs. SAMs traditionally use subhalos to track and determine galaxy properties and require specific subhalo information that is supplied from subhalo merger trees. The information inputs needed are the positions, velocities, masses, and evolution history of the subhalos, as well as a range of measurements for the infalling halo and the host halo that contains the subhalo. In order to use cores as a substitute for subhalos, we need to provide the analogs of these quantities for cores. Positions and velocities of the cores and properties of the

infall and host halos are already described by the core merger trees by construction. SMACC provides the additional mass information needed to enable galaxy modeling using SAMs.

This paper is the second in a series of papers focused on the Last Journey simulation. The simulation, described in the first paper (Heitmann et al. 2021), is a large, gravity-only run performed with the Hardware/Hybrid Accelerated Cosmology Code (HACC; Habib et al. 2016) and designed to enable the creation of detailed synthetic sky catalogs at different wavelengths. In the current paper, we focus on expanding the Last Journey core catalogs to generate a suitable set of inputs for SAM methods using SMACC. We investigate the robustness of subhalo mass modeling via SMACC as a function of mass resolution, the host halo concentration, individual subhalo-to-core matching, and also variations in cosmological parameters. We find that the underlying mass-loss model provides a very satisfactory description in all cases studied.

This paper is organized as follows. We begin in Section 2 with a brief overview of the simulations used in this paper. In Section 3, we review the definition of cores and describe how merger trees are used to enable core tracking. We demonstrate in Section 4 that a simple two-parameter mass-loss model for cores can be used to accurately predict the SHMF by comparing to results obtained with a subhalo finder. In Section 5, we discuss a range of measurements obtained with SMACC to verify the robustness of our approach. We present the results for our extreme-scale simulation, the Last Journey, in Section 6, and provide a summary of the results and future outlook in Section 7. In Appendix A we provide a short description of our subhalo finder and show that it yields results in good agreement with current literature. Finally, in Appendix B we present a short discussion about results at high redshifts.

## 2. Simulations

In this section, we summarize the simulations utilized in our analysis. The ultimate goal is to enable efficient substructure modeling for very large simulations, in particular the Last Journey simulation presented in Heitmann et al. (2021). For development and validation, however, smaller simulations are often very useful since they allow for faster data processing and testing. In addition, the analysis overhead does not require very large computing allocations or memory footprints. Consequently, we used a suite of smaller-scale simulations, described below, to carry out a set of modeling and convergence studies. All of these simulations were performed with the  $N$ -body code HACC, described in Habib et al. (2016). HACC has been designed to run at scale on a range of architectures; the simulations used in this paper have been run on IBM BG/Q and GPU-accelerated systems. We will also briefly describe the subhalo finder employed for comparisons to our new core-tracking methodology and show how well it matches results from other subhalo finders commonly used in the literature. Subhalo finding is carried out for our smaller simulations at a selected number of redshifts.

The three main simulations used in the paper are based on a Planck-18 cosmology (Planck Collaboration et al. 2020) with  $\theta = \{\Omega_{\text{cdm}} = 0.26067, \omega_b = 0.02242, h = 0.6766, \sigma_8 = 0.8102, n_s = 0.9665, w = -1\}$ , assuming a flat universe and massless neutrinos. The final target simulation is the Last Journey run, which evolved  $10,752^3$  particles in a  $(3400 h^{-1} \text{ Mpc})^3$  volume. The mass resolution for this simulation is  $m_p = 2.7 \cdot 10^9 h^{-1} M_\odot$ .

**Table 1**  
Simulations Used in This Paper

Name	Simulation Volume ( $h^{-1}$ Mpc) <sup>3</sup>	Particle (counts)	Particle Mass ( $h^{-1} M_{\odot}$ )	Core Host Halo Mass ( $h^{-1} M_{\odot}$ )	Core Size	Host Halo Mass ( $h^{-1} M_{\odot}$ )	Subhalo Mass ( $h^{-1} M_{\odot}$ )
Last Journey	3400 <sup>3</sup>	10,752 <sup>3</sup>	$\sim 2.7 \cdot 10^9$	$> 2.2 \cdot 10^{11}$ (80)	50	N/A	N/A
Last Journey-SV	250 <sup>3</sup>	1024 <sup>3</sup>	$\sim 1.3 \cdot 10^9$	$> 2.5 \cdot 10^{10}$ (20)	20	$> 8.8 \cdot 10^{10}$ (70)	$> 2.5 \cdot 10^{10}$ (20)
Last Journey-HM	250 <sup>3</sup>	3072 <sup>3</sup>	$\sim 4.6 \cdot 10^7$	$> 3.7 \cdot 10^9$ (80)	50	$> 9.3 \cdot 10^{10}$ (2,000)	$> 9.3 \cdot 10^8$ (20)
AlphaQ	256 <sup>3</sup>	1024 <sup>3</sup>	$\sim 1.2 \cdot 10^9$	$> 2.4 \cdot 10^{10}$ (20)	20	$> 8.4 \cdot 10^{10}$ (70)	$> 2.4 \cdot 10^{10}$ (20)

**Note.** Summary of the specifications of the four simulations used in this paper. The first four columns specify the simulations and their properties (simulation identifier, volume, particle count, and particle mass). The next four columns list the range of core host halo masses (the minimum halo mass above which a core is associated with a halo), the core size (number of core particles), the range of host halo masses (halos hosting at least one subhalo), and the range of subhalo masses. The numbers in parentheses for the reported mass thresholds are the corresponding number of simulation particles; the host halo mass is an FOF mass with  $b = 0.168$ . The force resolution for the three lower mass resolution simulations is  $\sim 2\text{--}3 h^{-1}$  kpc, and for the HM simulation it is  $\sim 0.8 h^{-1}$  kpc. The three “Last Journey” simulations are based on Planck-18 best-fit parameters, while the AlphaQ simulation is based on WMAP-7 best-fit parameters. (For more details, see text.)

Two smaller simulations were utilized for validation experiments. For the first run, we simply reduced the volume but kept the mass resolution similar to that in the Last Journey simulation. We chose a  $(250 h^{-1} \text{ Mpc})^3$  volume and  $1024^3$  simulation particles, leading to a mass resolution of  $m_p = 1.252 \times 10^9 h^{-1} M_{\odot}$ . The resulting smaller data set is much easier to handle and lends itself to a wide range of tests. In this case, to facilitate resolution measurements, we also reduce (to 20) the number of particles required to begin tracking cores (see Table 1, discussed later). We will refer to this simulation in the following as Last Journey-SV, or simply SV, for “small volume.” For the second run, we generated a simulation with much higher mass resolution, using the same volume of  $(250 h^{-1} \text{ Mpc})^3$  but increasing the number of particles to  $3072^3$ , leading to a mass resolution of  $m_p = 4.637 \times 10^7 h^{-1} M_{\odot}$ . The main purpose of this simulation is to enable detailed mass resolution and convergence studies for halo substructure. We will refer to this simulation in the following as Last Journey-HM, or simply HM, for high mass resolution.

Finally, to provide a variation in cosmology, we employ an additional simulation based on WMAP-7 results (Komatsu et al. 2011) in a  $(256 h^{-1} \text{ Mpc})^3$  volume, the AlphaQ simulation. The cosmological parameters used for the AlphaQ run are  $\theta = \{\Omega_{\text{cdm}} = 0.220, \omega_b = 0.02258, h = 0.71, \sigma_8 = 0.8, n_s = 0.963, w = -1\}$ . As for the Last Journey simulations, we assume flatness and massless neutrinos. The AlphaQ simulation is very similar to the Last Journey-SV simulation with regard to volume and mass resolution. The simulation was used in the past to develop a workflow to generate the synthetic galaxy catalog as part of the cosmoDC2 effort described in Korytov et al. (2019).

We summarize the simulation suite used in the paper in Table 1. In addition to the simulation specifications, we also provide information about minimum host halo and substructure masses for all simulations. For each simulation we find halos at 101 time steps between  $z = 10$  and  $z = 0$ , evenly spaced in  $\log a$  intervals (whenever we use  $\log$  it stands for  $\log_{10}$ ). For each snapshot, we store a halo catalog with detailed information covering a range of halo properties, including positions, velocities, shapes, and masses. A detailed description of the analysis performed for the Last Journey simulation is given in Heitmann et al. (2021).

### 3. Core Tracking

We have recently introduced the concept of core tracking in cosmological simulations as an alternative to the construction

of subhalo merger trees. A detailed description of the core catalog implementation is reported in Rangel et al. (2016) and Heitmann et al. (2021). The first scientific application of core tracking, in addition to comprehensive validation results, is discussed in D. Korytov et al. (2021, in preparation). In this section, we provide a brief summary of the underlying concepts and construction techniques for the core catalogs.

#### 3.1. Definition of Cores

The definition of a core is straightforward: For each halo above a certain mass (the core host halo mass threshold), the gravitational potential center is found and the set of particles closest to the center is identified as the halo core. The number of particles that define a core is an input parameter to the analysis setup and is chosen with regard to the size of the substructure that we aim to capture and the mass resolution of the simulation. The values for the core host halo mass threshold and the core size for our simulations are specified in Table 1. For each of the 101 analysis time snapshots, all core particles identified are stored, including their halo tag, to enable the connection to the halo catalogs at a later stage.

In addition to the core particle files per snapshot, we also keep track of the evolution of the core particles over time. Once a particle has been identified as a core particle, we follow its path for the remainder of the simulation. In totality, at each analysis step, we record the new positions and velocities of core particles that have been found previously and add newly found core particles as we go along. In this way, we build up an accumulated core particle file over time containing information about each core particle’s evolution from its first identification until the final time step. Storing the complete core history is necessary to generate core catalogs (as described in Section 3.3), which track the complex formation history of identified substructure.

#### 3.2. Merger Trees

In order to extract the temporal evolution of the substructure traced by core particles, halo merger trees are required. Such trees track the hierarchical formation history of each halo over time, recording merging events and mass accretion. The merger trees are constructed by processing halo catalogs at adjacent analysis snapshots. We compute the overlap of halo particles in each snapshot pair to connect older *progenitor* halos with younger *descendant* halos at the later time step. Processing all snapshots results in a complete merger tree catalog that can extract the formation history of any halo of interest.



There are many complications in merger tree construction, with a number of papers devoted to the subject (e.g., Fakhouri & Ma 2008; Behroozi et al. 2013; Rodriguez-Gomez et al. 2015; Han et al. 2018). Finite mass resolution, for example, can cause stochastic threshold effects of the smallest halos, which are repeatedly found, lost, and refound at masses near the halo minimum mass threshold. We have mitigated this complication by setting our halo finding thresholds to be below our minimum host halo masses in our trees, effectively pruning out halos of masses below our tracking resolution. The smallest resolved halos in our merger trees are recorded by the core host halo mass in Table 1.

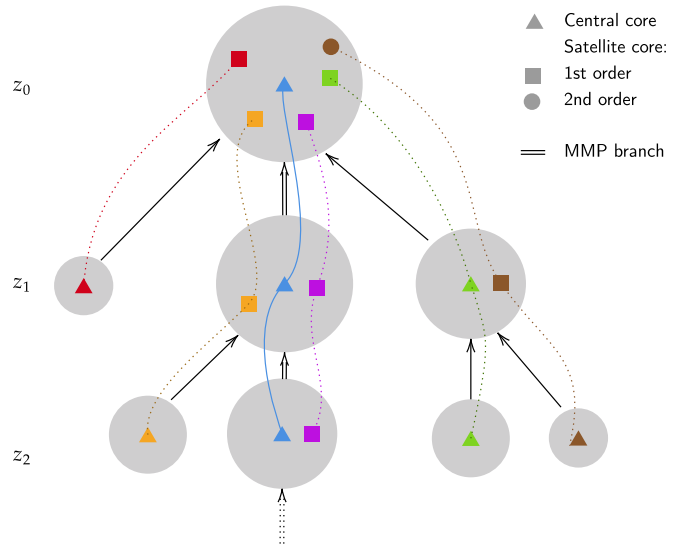
Merger tree construction is further complicated by halo “splitting” events, where nearby (or flyby) halos are temporarily considered as one object (overlinked by a friends-of-friends (FOF) finder, for example), erroneously indicating a merger event, and are later discovered to be different separate (split) objects. To avoid splitting effects, we have found that the construction of merger trees is best carried out in post-processing, starting from the final time step of interest (in our case at  $z=0$ ) and walking backward in time. Thus, we are ensured that every merger tree will be rooted by one final descendant, and we employ an artificial halo “fragmentation” procedure to account for halo splitting events. Briefly, when we detect a halo progenitor with multiple descendants (indicating the halo split), we fragment the progenitor into separate halos defined by the overlapping particles of each descendant. We then assign to each fragment a portion of the original progenitor mass based proportionally on the descendant mass, thereby conserving total mass (see Rangel et al. 2016 for details). The resulting merger trees are devoid of splitting events and consistently track the formation history of each halo.

In summary, we efficiently record the merger history of all halos to form a connectivity tree catalog that consistently tracks all mass that eventually ends up in the final halos of interest at  $z=0$ . For each object within the tree, we further store all of the relevant halo catalog properties (mass, velocity, shape, etc.), with adjusted values used for fragmented halos when needed.

### 3.3. Core Catalogs

By combining the information from core tracking, halo properties, and merger trees, one can construct core catalogs that contain the evolution history of each core from its birth to the final time step. Core catalogs are analogous to subhalo merger trees and can track the substructure evolution of halos. A detailed description of core catalog construction is given in D. Korytov et al. (2021, in preparation). Figure 1 shows an example of a halo merger tree that also includes core information. Each halo in the merger tree contains one central core, as well as satellite cores that are inherited by merged progenitor halos.

Core catalogs are constructed by traversing merger trees forward in time. For every halo in the tree, we extract the central particles to determine its central core. When halos merge, the central core of the most massive progenitor (MMP) is connected to the measured central core of the descendant halo. The other less massive progenitors contribute their own central cores to the descendant, which in turn are labeled as satellite cores. The satellites will trace the dynamical substructure evolution of these progenitor halos. Moreover, if any progenitor contains satellites from previous mergers, these objects are also carried into the descendant halo, maintaining the hierarchical information of previous structure formation.



**Figure 1.** Example halo merger tree showing the types of cores associated with halos (gray disks). Central cores (triangles) are identified when an FOF halo is found above a preset core host halo mass threshold. As halos merge, the central core of the MMP is connected to the measured central core of the descendant. Cores of less massive progenitors are designated as satellites. The “order” of a satellite traces the number of previous mergers away from an MMP; cores that merged directly into an MMP halo are labeled first-order satellites (squares), cores inherited from one generation above an MMP merger are second-order satellites (circles), etc. This ordering of cores is analogous to substructure hierarchies described in the literature (e.g., Giocoli et al. 2010; Jiang & van den Bosch 2014, 2016).

Analogous to the subhalo hierarchies typical of those described in the literature (e.g., Giocoli et al. 2010; Jiang & van den Bosch 2014, 2016), one can assign a satellite “order” based on the generation in which a central core was merged and converted to a satellite. Central cores that merged directly into an MMP are considered *first-order* satellites, central cores merging into halos that later merge into an MMP themselves are labeled *second-order*, and so on, as illustrated in Figure 1.

As all core particles are stored in the simulation output data since birth, we can accurately update satellite trajectories within halos by extracting the updated core particle positions and velocities at each snapshot. Moreover, every satellite core carries the infall mass and redshift of its host halo when it merged. These quantities will inform the mass-loss model for the cores that we discuss in Section 4.

## 4. Mass Modeling

Developing and verifying a viable mass-loss model for core-based substructure is a necessary step in the creation of a usable proxy for subhalo merger trees. As described in Section 3.3, satellite cores are substructure tracers of merged halos. We aim to model the mass loss of a satellite object during its post-merger evolution. In this section, we describe our mass-loss model implementation, compare our results to subhalo finder measurements to determine appropriate fitting parameters, and report on resolution tests.

The primary metric we use to assess the accuracy of the mass-loss modeling approach is the SHMF, given by  $dn/d\log(m/M)(z)$ , which is a measure of the number of subhalos (of mass  $m$ ) within a host halo (of FOF mass  $M^6$ ) at

<sup>6</sup> We use the FOF mass for binning, not the fragment mass (defined in Section 3.2), to be consistent with the subhalo finder measurements that are recorded using FOF mass.

some redshift  $z$ . Unless otherwise specified, the SHMF bins are normalized by the total number of hosts, as well as the bin width  $d \log(m/M)$ . The reported error bars are determined by the statistical error of each bin. We compare the SHMF of our mass-evolved satellite cores with the corresponding mass function generated from our subhalo finding implementation described in Appendix A.

#### 4.1. Subhalo Mass Modeling

The approach we choose to follow was introduced by van den Bosch et al. (2005) and uses a power law to represent the average subhalo mass-loss rate. Their model was motivated by the study of idealized isolated circular subhalo orbits. The mass-loss rates of subhalos in  $N$ -body simulations have been measured by, e.g., Giocoli et al. (2008), who found good agreement with the power-law model proposed by van den Bosch et al. (2005). Additional studies have extended the model (Jiang & van den Bosch 2016; van den Bosch & Jiang 2016; Hiroshima et al. 2018). We begin by summarizing the mass-loss equations, in a manner consistent with the notation conventions of Jiang & van den Bosch (2016).

The average subhalo mass-loss rate over all orbital configurations is assumed to be given by

$$\dot{m} = -\mathcal{A} \frac{m}{\tau_{\text{dyn}}} \left( \frac{m}{M} \right)^{\zeta}, \quad (1)$$

where  $M = M(z)$  and  $m = m(z)$  are the parent halo mass and subhalo mass at redshift  $z$ , respectively, and  $\mathcal{A}$  and  $\zeta$  are free parameters. The dynamical time of the halo is

$$\tau_{\text{dyn}}(z) = 1.628 h^{-1} \text{Gyr} \left[ \frac{\Delta_{\text{vir}}(z)}{\Delta_{\text{vir}}(0)} \right]^{-1/2} \left[ \frac{H(z)}{H_0} \right]^{-1}, \quad (2)$$

where  $\Delta_{\text{vir}}(z)$  is the virial overdensity. Following van den Bosch et al. (2005) and Jiang & van den Bosch (2016), we use the fitting function of Bryan & Norman (1998) for the virial overdensity:  $\Delta_{\text{vir}}(z) = 18\pi^2 + 82x - 39x^2$ , where  $x(z) = \Omega_m(z) - 1$  and  $\Omega_R = 0$ .

If the host halo mass is constant ( $\dot{M} = 0$ ),

$$m(t + \Delta t) = \begin{cases} m(t) \exp(-\Delta t / \tau) & \zeta = 0, \\ m(t) \left[ 1 + \zeta \left( \frac{m}{M} \right)^{\zeta} \frac{\Delta t}{\tau} \right]^{-1/\zeta} & \text{otherwise.} \end{cases} \quad (3)$$

The timescale,  $\tau(z)$ , is the characteristic timescale of subhalo mass loss and is defined as  $\tau(z) \equiv \tau_{\text{dyn}} / \mathcal{A}$ . The mass-loss model was originally developed to process EPS merger trees (see Zentner 2007 for a review of EPS), where the two free parameters were selected to best fit the average SHMF of two  $\Lambda$ CDM simulations for a host halo mass of  $M = 10^{15} h^{-1} M_{\odot}$ . The decay model was found to be in reasonable agreement with the simulation measurements and has been studied and refined in a number of follow-up papers (e.g., Jiang & van den Bosch 2016). We have found that the simple two-parameter model is well suited for our goals and provide some of the implementation details in the following.

For our core mass-loss model, we implement Equation (3) using core catalogs in place of EPS merger trees. The satellites in the core catalog carry the host halo merger tree information required for the mass-loss model implementation. For each

satellite core, we store the simulation time step immediately preceding infall (merging) of the former host halo, labeled  $t_{\text{infall}}$ , in addition to the halo properties at  $t_{\text{infall}}$  (including mass), which we call the satellite’s infall properties. We then initialize the model based on the infall mass and evolve the core mass at each subsequent snapshot using Equation (3) ( $\tau$  is computed at the redshift corresponding to  $t + \Delta t$ ). Note that we use “core mass” to refer to the modeled mass of the substructure object associated with the core’s position.

The parent mass  $M(z)$  used to calculate the mass evolution of a core in Equation (3) is chosen to be the mass of its immediate parent halo (potentially traced by another core with its own evolved mass). Explicitly, for cores of order  $>1$  (as defined in Section 3.3),  $M(z)$  is the evolved mass of the parent halo into which the core first fell. For cores with order  $=1$ , the parent halo mass is simply the host halo mass measured at the given time step; note that we specifically use the fragment host halo mass defined in Section 3.2, to be consistent with our merger tree construction. This mass integration approach is similar to the procedure of Jiang & van den Bosch (2016) for evolving the mass of subhalos that can have order greater than 1.

A brief remark on initialization for the mass model: for objects that are identified at snapshot  $t_{\text{infall}}$  and have merged into a host halo by the next snapshot, labeled  $t_{\text{satellite}}$ , the merging event could have occurred at any time between  $t_{\text{infall}}$  and  $t_{\text{satellite}}$ . We begin the mass evolution at the average of the two snapshots  $(t_{\text{infall}} + t_{\text{satellite}})/2$ . Thus, the evolved mass of the core at  $t_{\text{satellite}}$  is given by Equation (3), with  $\Delta t = (t_{\text{satellite}} - t_{\text{infall}})/2$ ,  $m$  given by the infall (fragment) mass, and  $M$  given by the host halo mass at  $t_{\text{satellite}}$ . The mass evolution for subsequent snapshots is calculated as described above.

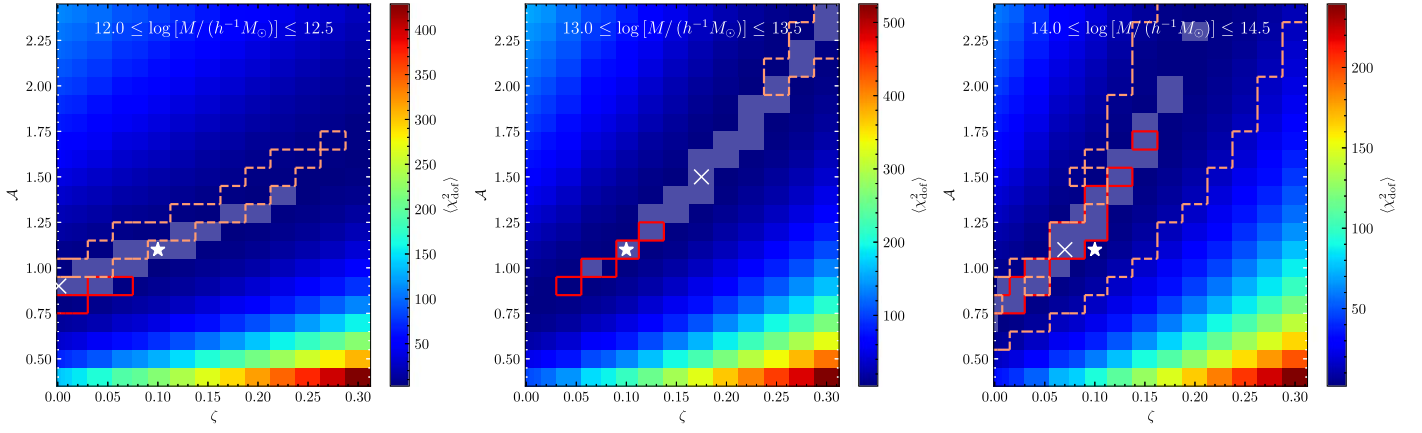
Next, we discuss our approach to find the appropriate fitting parameters for the mass-loss model ( $\mathcal{A}$  and  $\zeta$ ) and evaluate the validity of the overall SMACC approach.

#### 4.2. Parameter Exploration

Given the mass-loss model introduced in Section 4.1, we need to determine optimal parameter values of  $\mathcal{A}$  and  $\zeta$  for Equation (3). As the simulations underlying our core catalogs differ from those of van den Bosch et al. (2005), we expect a different set of optimal model coefficients. We choose as our fitting criteria the comparison of the predicted SHMF from the core catalogs to that produced by a subhalo finder (see Appendix A). We use our HM simulation results at  $z = 0$  and  $z = 1$  for this analysis. As shown in Figure 10 for  $z = 0$ , the HM subhalo finder results agree with the results of SV in the higher subhalo mass overlap region, while the higher mass resolution of HM allows an exploration of lower-mass subhalos past the resolution limits of SV.

For a given host halo mass bin, we compute the average SHMF  $dn/d \log(m)$  within  $N_b = 20$  bins of equal width in a specified  $\log(m)$  range. Note that we do not normalize  $m$  by  $M$  for this analysis. We have run our mass model over a grid of the free parameters, spanning

$$\begin{aligned} \mathcal{A} &= \{0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1.0, 1.1, 1.2, 1.3, 1.4, 1.5, \\ &\quad 1.6, 1.7, 1.8, 1.9, 2.0, 2.1, 2.2, 2.3, 2.4\}, \\ \zeta &= \{0.001, 0.005, 0.01, 0.02, 0.04, 0.07, 0.08, 0.1, 0.125, \\ &\quad 0.15, 0.175, 0.2, 0.225, 0.25, 0.275, 0.3\}. \end{aligned}$$



**Figure 2.** Measurements of  $\langle \chi_{\text{dof}}^2 \rangle$ , i.e., the mean of  $\chi_{\text{dof}}^2$  over  $z = 0$  and  $z = 1$ ; all results are for the HM simulation. We limit our fitting to cores and subhalos with  $m > 100m_{\text{p,SV}}$  (where  $m_{\text{p,SV}}$  is the particle mass of the SV simulation) and show the  $(\mathcal{A}, \zeta)$  free parameter exploration for three different host halo mass bins. For each host mass bin, the region in parameter space where  $\Delta\langle \chi_{\text{dof}}^2 \rangle \leq 1$  is shaded in white, and the best-fit parameters are marked with a cross. We also outline the region where  $\Delta\chi_{\text{dof}}^2(z=0) \leq 1$  (red solid line) and the region where  $\Delta\chi_{\text{dof}}^2(z=1) \leq 1$  (pink dashed line). The mass regions over which we calculate  $\langle \chi_{\text{dof}}^2 \rangle$  for the three host mass bins and the resulting best-fit parameters are recorded in Table 2. The star indicates our fiducial model parameters ( $\mathcal{A}, \zeta = 1.1, 0.1$ ) that we use in the analyses that succeed Section 4.2.

**Table 2**  
Host Halo Mass Bins for Model Fitting (HM)

$\log[M/(h^{-1}M_\odot)]$	Host Halo Count	Fitting Range in $\log(m)$	$(\mathcal{A}, \zeta)_{\min(\langle \chi_{\text{dof}}^2 \rangle)}$	$\min(\langle \chi_{\text{dof}}^2 \rangle)$
[12.0, 12.5]	39,505 (33,020)	$(\log(100m_{\text{p,SV}}), 11.9)$	(0.9, 0.001)	2.70
[13.0, 13.5]	4738 (2368)	$(\log(100m_{\text{p,SV}}), 12.5)$	(1.5, 0.175)	3.41
[14.0, 14.5]	348 (21)	$(\log(100m_{\text{p,SV}}), 13.0)$	(1.1, 0.07)	1.77

**Note.** Mass bins listed in this table are used to optimize the mass-loss modeling parameters. Main values of the second column correspond to  $z = 0$ , and the values in parentheses correspond to  $z = 1$ .  $m_{\text{p,SV}}$  is the particle mass of the SV simulation. Results are shown in Figures 2 and 3. The three bins are also used for the verification test in Figure 4.

This allows us to compute the average core-evolved mass function  $dn/d\log(m)$  for the same bins at any point of the grid,  $\text{CMF}_i(\mathcal{A}, \zeta)$ . We define

$$\chi_{\text{dof}}^2(\mathcal{A}, \zeta) = \frac{\chi^2(\mathcal{A}, \zeta)}{N_b - N_p}, \quad (4)$$

where

$$\chi^2(\mathcal{A}, \zeta) = \sum_i \left[ \frac{\text{SHMF}_i - \text{CMF}_i(\mathcal{A}, \zeta)}{\sigma_i(\mathcal{A}, \zeta)} \right]^2, \quad (5)$$

$$\sigma_i(\mathcal{A}, \zeta) = \sqrt{\Delta(\text{SHMF}_i)^2 + \Delta(\text{CMF}_i(\mathcal{A}, \zeta))^2}. \quad (6)$$

Here  $\text{SHMF}_i$  and  $\text{CMF}_i$  are the SHMFs and core-evolved mass functions in bin  $i$ , respectively, and  $\Delta(\text{SHMF}_i)$  and  $\Delta(\text{CMF}_i)$  are the statistical errors for each bin. The number of fitting parameters  $N_p = 2$ . Finally, we compute  $\chi_{\text{dof}}^2$  over the grid scan of the parameters at  $z = 0$  and  $z = 1$  independently and measure the mean over the two redshifts,

$$\langle \chi_{\text{dof}}^2 \rangle = \frac{1}{2} [\chi_{\text{dof}}^2(z=0) + \chi_{\text{dof}}^2(z=1)]. \quad (7)$$

The resulting values of  $\langle \chi_{\text{dof}}^2 \rangle$  are shown in Figure 2 for three host halo mass bins of the Last Journey-HM simulation. Table 2 shows the host halo mass ranges and counts for the three bins, as well as the ranges of  $\log(m)$  over which we calculate  $\chi_{\text{dof}}^2$  for each redshift. Note that we use a mass of 100 particles of the SV simulation ( $\log m \sim 11.1$ ) as the fitting range lower bound for all three mass bins; we show in

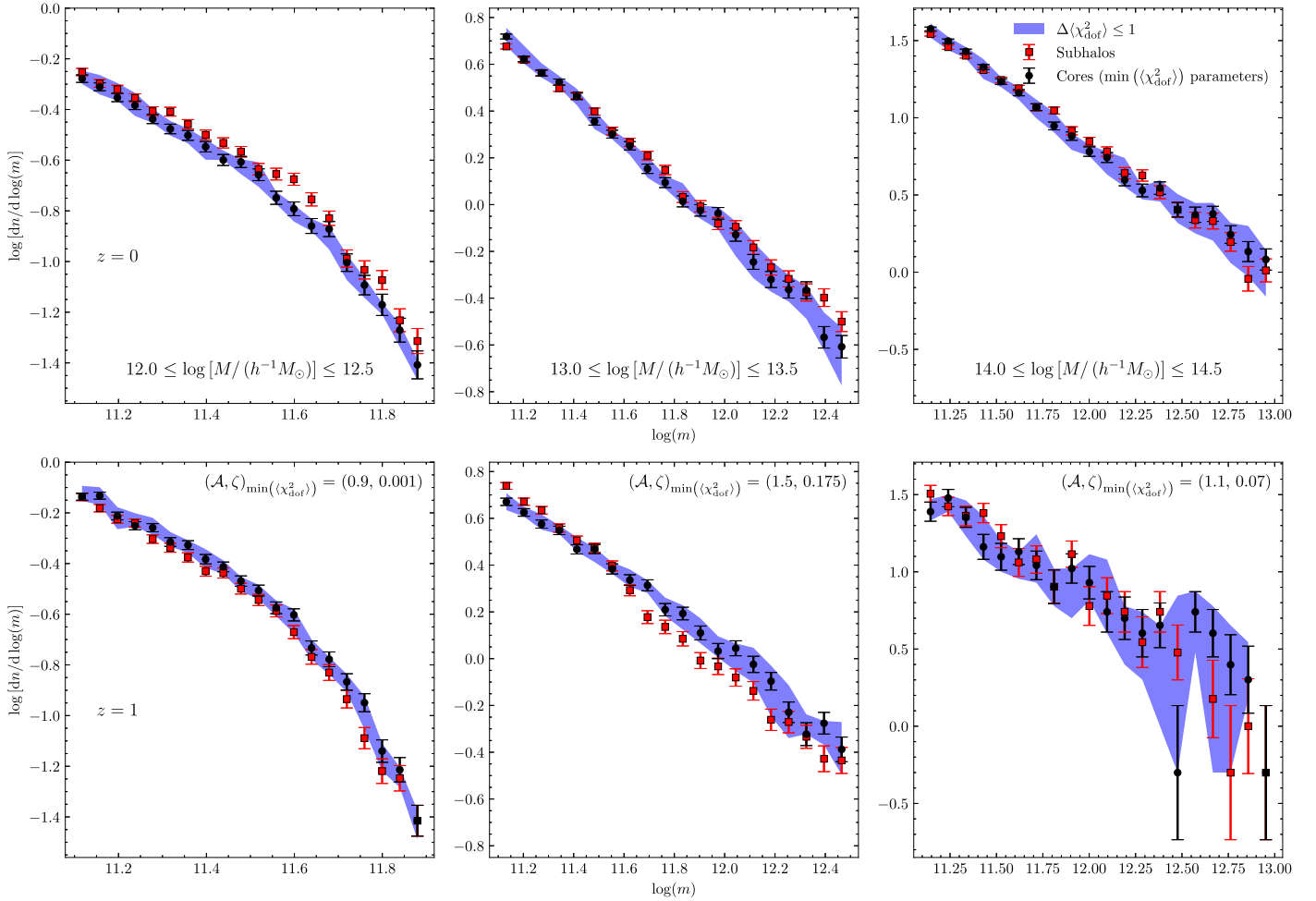
Section 4.3 that this threshold is the approximate resolution limit above which the HM and SV substructure mass functions are mutually converged. The fitting range upper bounds were chosen to be just before the high subhalo masses fall off of the SHMF.

For each host bin, we designate the grid point where  $\langle \chi_{\text{dof}}^2 \rangle$  is minimized to be the best-fit parameters. This point in parameter space is marked by a cross in Figure 2. Table 2 shows the free parameter values and  $\langle \chi_{\text{dof}}^2 \rangle$  at this point. The light shaded region in each panel of Figure 2 indicates the grid points for which  $\Delta\langle \chi_{\text{dof}}^2 \rangle \leq 1$ , where

$$\Delta\langle \chi_{\text{dof}}^2 \rangle(\mathcal{A}, \zeta) = |\langle \chi_{\text{dof}}^2 \rangle(\mathcal{A}, \zeta) - \min(\langle \chi_{\text{dof}}^2 \rangle)|. \quad (8)$$

We also indicate the individual best-fit regions of each redshift in the three panels: the region where  $\Delta\chi_{\text{dof}}^2(z=0) \leq 1$  is outlined by the red solid line, and the region where  $\Delta\chi_{\text{dof}}^2(z=1) \leq 1$  is outlined by the pink dashed line. Note that  $\Delta\chi_{\text{dof}}^2(z)$  for any point in parameter space is defined by Equation (8) if  $\chi_{\text{dof}}^2(z)$  is substituted for  $\langle \chi_{\text{dof}}^2 \rangle$ .

In Figure 3 we show the (shaded) range of the core mass functions at  $z = 0$  (top panels) and  $z = 1$  (bottom panels) for model parameters measuring  $\Delta\langle \chi_{\text{dof}}^2 \rangle \leq 1$  for the three host mass bins of Table 2. We also show the SHMF and best-fit core mass function in the  $\log(m)$  fitting range. First, we note that these two functions are in good agreement across all three mass bins shown. Second, the range of allowed models for the two-parameter fit is rather narrow, establishing that the mass-loss



**Figure 3.** Comparison of the subhalo and core mass functions. For three host halo mass bins at  $z = 0$  (top panels) and  $z = 1$  (bottom panels) for HM, the shaded region shows the range of core mass functions with the mass model parameters for which  $\Delta\langle\chi^2_{\text{dof}}\rangle \leq 1$ . Also shown is the SHMF and the core mass function using the best-fit (minimum  $\langle\chi^2_{\text{dof}}\rangle$ ) model parameters found for each bin (parameter values are indicated in the bottom panels). All mass functions are average mass functions computed in the same mass range as the parameter fitting. See Table 2 for the best-fit parameter values, the mass ranges used for fitting, and a listing of the statistics for each bin. The best-fit mass-loss model reproduces the SHMF well; we note that there is a slight redshift dependence (exemplified by the small overall offsets between the model and subhalo measurements in the low-mass bin) that is not fully captured by the simplistic model; however, the deviations are small. Note that while we show the particular best-fit parameters of each mass bin in this figure, we use the fiducial model parameters ( $\mathcal{A}, \zeta = 1.1, 0.1$ ) in the analyses that follow Section 4.2.

formulation provides a well-constrained model for matching cores to subhalos.

By comparing our subhalo and core model results at  $z = 0$  and  $z = 1$ , we have found that the core mass function best-fit parameters are different for the two redshifts. This indicates a possible redshift dependence in the mass modeling parameters. The redshift deviation is further demonstrated when comparing the mass functions in the lowest-mass bin of Figure 3; the best-fit core mass function is biased low compared to the SHMF at  $z = 0$ , whereas the opposite is true at  $z = 1$ . This is due to our chi-square optimization over the average of the two redshifts. Additionally, the varying best-fit parameter values across the host mass bins shown in Table 2 imply a host halo mass dependence of the fitting parameters as well.

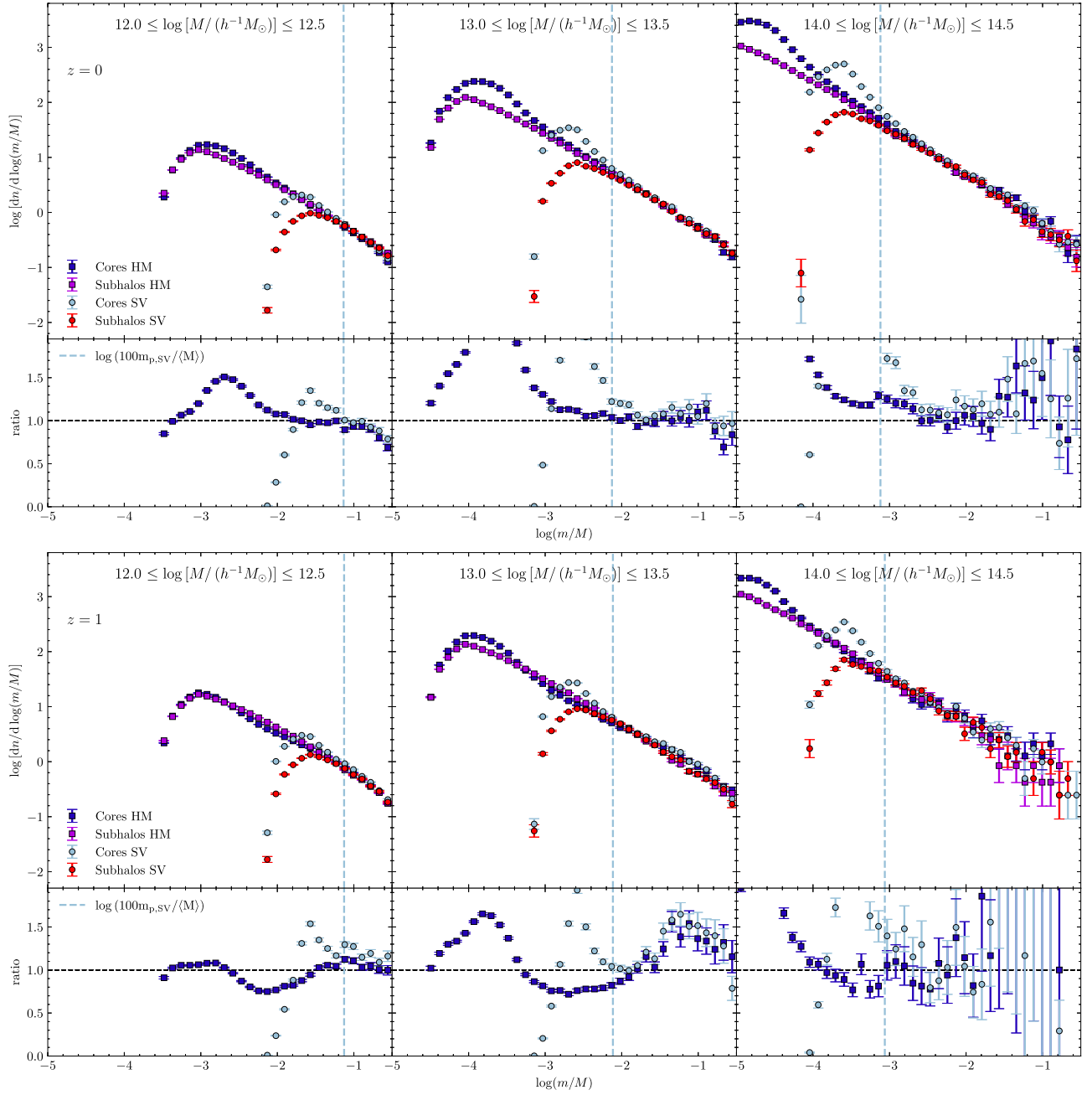
For simplicity, we choose  $(\mathcal{A}, \zeta) = (1.1, 0.1)$  as our fiducial modeling parameters to use across all redshifts and host halo masses for which we compute the mass model. As shown in Figure 2, this choice of parameters is within the  $\Delta\langle\chi^2_{\text{dof}}\rangle \leq 1$  region for the first two mass bins and within the  $\Delta\chi^2_{\text{dof}}(z = 1) \leq 1$  region for the high-mass bin. Despite the redshift and host halo mass dependence of the model parameters, we show in

Figure 4 that, using the fixed parameters, we are able to produce substructure mass functions at  $z = 0$  and  $z = 1$  that have good agreement with the subhalo finder results for both HM and SV to our resolution tolerance (see Section 4.3). These parameters provide a reasonable compromise across the two redshifts and three mass bins examined in this section using HM. (We have verified this result for the SV simulation by carrying out a similar analysis.) The fixed parameters are indicated in Figure 2 by a star. We will use the fiducial parameters  $(\mathcal{A}, \zeta) = (1.1, 0.1)$  for all other comparisons. We provide a brief discussion regarding the performance of the mass-loss model at higher redshifts in Appendix B.

#### 4.3. Resolution Tests

Measuring model robustness to significant changes in simulation mass resolutions is important to understand the numerical limitations and tolerances of SMACC. We expect resolution effects on our core mass model, as the particle sets that trace substructure, in addition to the parent halos themselves, are susceptible to the force and mass resolution limitations of the





**Figure 4.** Comparison of the mass model applied to two simulations with different mass resolutions to test resolution effects. Top panels: average subhalo and core mass functions at  $z = 0$  for three host halo mass bins in the HM (squares) and SV (circles) simulations. Subhalos and cores with  $m > 20$  particles are shown. The HM host bins contain 39,505, 4738, and 348 host halos, while the SV bins contain 40,628, 4797, and 337 halos. The vertical dashed line in each panel approximates the 100-particle mass threshold for SV, above which our model predictions appear to be converged. We additionally measure the ratio of the core mass functions to the HM SHMF, displayed at the bottom of each panel. Bottom panels: same as the top panels, but for  $z = 1$ . In this case the HM host bins contain 33,020, 2368, and 21 host halos, while the SV bins contain 34,592, 2418, and 36 halos.

underlying simulation. Moreover, we expect that our core mass model will have a resolution-dependent lower mass limit; above such a threshold, our results would be converged and represent the minimum mass of substructures we can reliably model. During our measurements, we want to ensure that the resolution mass cuts we employ for the core catalog in the large Last Journey simulation result in a complete set of halo cores equivalent to subhalos measured above the same mass threshold. We remind the reader that a major aim of the Last Journey series is to generate a set of core merger trees that can be used as an alternative to subhalo merger trees as input to semianalytic models.

In order to investigate the resolution effects, we use the core catalogs from the Last Journey-SV and HM simulations described in Section 2, at redshifts  $z = 0$  and  $z = 1$ . (The mass resolution in the two simulations differs by almost a factor of 30.) We also employ the results from the subhalo finder for this test. Figure 4 shows our results for three host halo mass bins (the first three bins listed in Table 2) at both redshifts. For this test, we identified subhalos with over 20 particles and tracked cores at the same resolution. Each of the three panels shows two sets of curves corresponding to the core and SHMFs from both simulations. We found a minimum core mass threshold of



$\sim 100$  particles of SV to be sufficient to limit resolution effects of the simulation in our mass model. We mark the mass cut in each panel via the vertical dashed line.

Collectively, for all host halo mass bins at both redshifts, the results in Figure 4 for the high- and low-resolution simulations demonstrate excellent agreement for the measured core and SHMFs up to the marked mass cut limit. This result is very satisfying, as it shows that our results are indeed converged. Moreover, the individual agreement between the subhalo and core mass functions for each simulation demonstrates the accuracy of our optimized mass model. As the parameter tuning discussed in Section 4.2 was carried out on the HM simulation, the matching results of the SV simulation utilizing the same parameters are an indication that our model fitting was robust. Lastly, we note that both simulations exhibit a fall-off in the mass functions after their individual resolution limits are reached.

As the Last Journey simulation ran with a similar mass resolution to SV, an equivalent threshold of 100 particles correspondingly produces accurate mass-evolved substructure measurements. The results of Figure 9 in Section 6 demonstrate the quality of the modeled SHMF in the Last Journey analysis using the reported mass cut. In summary, the resolution tests show the robustness of our approach to mass resolution and allow us to confirm that our chosen mass limit for the final SAM project is valid.

## 5. Model Robustness

In this section we present results from further investigations of the robustness of the SMACC approach. We compare predictions from SMACC directly with measurements from our subhalo finder, going beyond the SHMF shown in the previous section. We focus on the analysis of the SV simulation since its mass resolution is representative of the main Last Journey simulation. We will discuss effects of the host halo concentration, in addition to results for a one-to-one match-up comparison with subhalos and their associated cores. Lastly, we investigate possible cosmology dependence on our chosen model parameters, making use of the AlphaQ simulation. All of these tests are performed on subhalos and cores with evolved masses above a threshold of 100 particles, determined in Section 4.3 to represent sufficiently resolved objects. Our findings provide additional evidence that SMACC enables the use of core catalogs to act as a substitute for subhalo merger trees.

### 5.1. Host Halo Concentration Effects

The SMACC approach makes the strong assumption that the core mass  $m$  and the ratio with the host halo mass  $m/M$  are the only ingredients needed to reliably predict the mass-loss evolution of substructures inside a host halo. However, structure formation is a complex process, and the environment in which a halo formed and grew, in addition to other halo properties such as formation time, surely influences substructure evolution. In order to capture some of this additional information, we employ the host halo concentration as a simple proxy for estimating these effects. The concentration does depend on the environment (see, e.g., Wechsler et al. 2002 for a detailed discussion) and therefore will provide a good indication for the robustness of our approach. For example, tidal forces will be stronger in higher-concentration halos, so

one would expect more subhalo mass loss in this population. The concentration also carries information about the halo’s evolution history and local environment. Within a given mass range, halos that have lived a quiet life will have higher concentrations. This implies, once again, a lower SHMF because the subhalo or core mass “feeding function” for these halos will likely be smaller.

To investigate the range of possible effects, we partition the halo population by sorting halos within a given mass range into two concentration groups. We restrict our investigation to a mass bin of  $10^{12} - 10^{13} h^{-1} M_{\odot}$  for the host halo. While the host halo mass is given by the FOF,  $b = 0.168$  mass, the concentration is measured using the spherically averaged density profile and the associated spherical overdensity mass,  $M_{200c}$ . For each FOF halo we determine the potential minimum and grow a sphere around it, measuring the radial density profile and fitting a Navarro–Frenck–White (NFW) functional form (Navarro et al. 1996, 1997) to it. We first determine the mass and then fit for the halo concentration. The NFW profile is given by

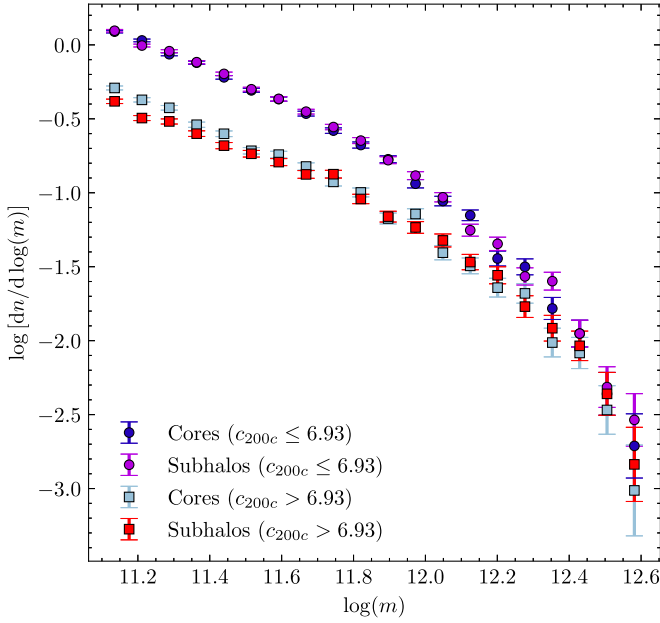
$$\rho(r) = \frac{\delta \rho_{\text{crit}}}{(r/r_s)(1 + r/r_s)^2}, \quad (9)$$

where  $\delta$  is a characteristic dimensionless density and  $r_s$  is the scale radius. The halo concentration is defined via  $c_{\Delta} = r_{\Delta}/r_s$ , where we take  $\Delta$  to be the overdensity with respect to the critical density of the universe,  $\rho_{\text{crit}} = 3H^2/8\pi G$ . The radius,  $r_{\Delta}$ , is reached when the enclosed mass,  $M_{\Delta}$ , equals the volume of the sphere times the density  $\Delta \rho_{\text{crit}}$ . We measure masses and concentrations for  $\Delta = 200$ , corresponding in turn to  $c_{200c} = R_{200c}/r_s$ . A more detailed discussion of our approach can be found in Child et al. (2018), where alternative methods for concentration measurements are discussed as well.

Figure 5 shows the substructure mass functions for host halos in two different concentration ranges of  $c_{200c} \leq \text{Med}(c_{200c})$  and  $c_{200c} > \text{Med}(c_{200c})$ , where the median concentration of the host halos in the mass bin is  $\text{Med}(c_{200c}) = 6.93$ . As in the previous tests, we show both the SHMF and the core mass function. We consider only subhalos and cores that have a mass over the resolution threshold of 100 particles. First, we note that the mass function for the high-concentration halos is increasingly suppressed, compared to that for the low-concentration halos, as the substructure mass decreases. This is due to the connection between concentration and halo age: the more concentrated halos are older, and therefore substructures have had more time to either merge into the central, high-density region or be disrupted. This finding is in agreement with previous results (see, e.g., Gao et al. 2004; Zentner et al. 2005; Wechsler et al. 2006). Second, the agreement between the evolved core masses and the subhalo masses in both concentration bins is still excellent. This result confirms that the SMACC approach is robust against different halo formation histories for modeled masses above our resolution threshold.

### 5.2. Subhalo–Core Spatial Matching

The mass functions used thus far for comparisons between subhalos and cores characterize agreement in overall mass distributions. We have shown that the core mass-loss methodology, in the sense of averages (e.g., SHMFs), performs well. The final step is to investigate and confirm that the simple



**Figure 5.** Average subhalo and core mass functions of the SV simulation at  $z = 0$  computed for two host halo concentration bins. Note that in this case the  $x$ -axis is the substructure mass (subhalo and modeled core masses) and that we only consider subhalos and cores that have a mass corresponding to over 100 particles. Host halos in the mass range  $12.0 \leq \log[M/(h^{-1}M_{\odot})] \leq 13.0$  are divided into two bins based on the halo concentration  $c_{200c}$ . The first half  $c_{200c} \leq \text{Med}(c_{200c})$  and second half  $c_{200c} > \text{Med}(c_{200c})$  bins contain 27,058 and 27,057 halos, respectively, where the host halo median concentration  $\text{Med}(c_{200c}) = 6.93$ . Note that our host halos have an FOF mass definition, while their concentrations are derived from the corresponding spherical overdensity halos with  $\Delta = 200$  (with respect to the critical density). About 1.5% of the 54,914 FOF halos within the mass range are removed from our analysis because they are not well fit by a spherical profile.

description of allowing cores and their history to be associated with a substructure mass is in fact valid at the level of individual objects. This is clearly a necessary requirement for the core approach to be used with detailed semianalytic models of galaxy formation.

We now investigate how well the mass model works on a case-by-case basis by comparing the mass of an individual subhalo to its corresponding modeled core. For this analysis, we use our SV simulation results at  $z = 0$  and consider only subhalos and cores that have a resolved mass over 100 particles. We perform a spatial matching procedure by selecting the most massive core within a search radius of  $2r_{\text{sh, vir}}$  of each subhalo. We define the virial radius of a subhalo,  $r_{\text{sh, vir}}$ , as the radius corresponding to a uniform sphere of mass equal to the subhalo mass and density equal to  $\Delta_{\text{vir}}(z)\rho_c(z)$ , where  $\rho_c(z)$  is the critical density (note that we consider the subhalo mass to be approximately a virial mass for this calculation). The subsets of subhalos that have at least one core (above the mass cut) within the search radius are 74%, 85%, and 90% of the total subhalos (above the mass cut) for three host halo mass bins covering a mass range of  $[12.0, 12.5]$ ,  $[13.0, 13.5]$ , and  $[14.0, 14.5]$  in  $\log[M/(h^{-1}M_{\odot})]$ , respectively.

Figure 6 shows the distribution of the relative difference in mass of the subhalo and matched core for the three host halo mass bins described above. We show results for the infall mass and the evolved core mass (using the  $z = 0$  subhalo mass for calculating both relative differences), with the expectation that the evolved core mass should show a significantly improved correspondence, i.e., the distribution should shift significantly

toward the origin with an increase in the peak height. This is indeed observed, and, in particular, for the smaller host halo mass bin, the agreement is rather good. For the larger host halo masses, the matching will be less precise since there are more subhalos per host halo that could provide an alternative (and more suitable) match. Nevertheless, even for the two higher-mass bins, the results show that most evolved core masses are close to the corresponding subhalo mass. In Figure 7 the mass comparison for the same data is shown in a different way to study possible biases in the matched masses. Here we show mass comparison contours for the same set of matched subhalos and cores and three mass bins as in Figure 6. The contours align very well on the diagonal, indicating good agreement (with some intrinsic scatter) for the matched subhalo-evolved core mass pairs.

### 5.3. Cosmology Effects

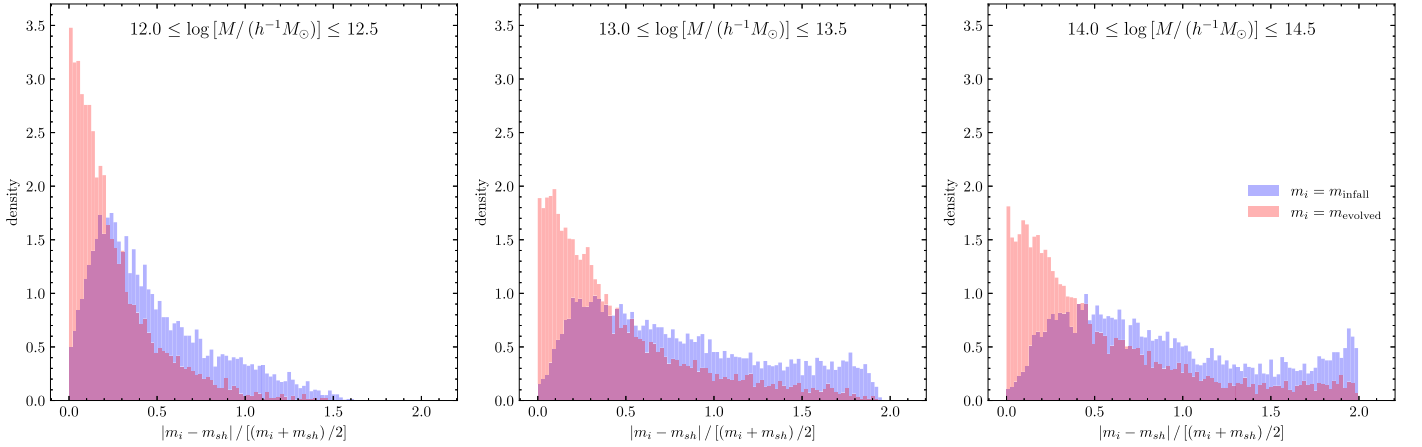
Finally, we investigate the sensitivity of the core mass modeling approach to a change in the cosmological parameters specifying the simulation. For this test, we employ two simulations: the first is the Last Journey-SV simulation, using the Planck-18 cosmology, while the second, the AlphaQ simulation, is based on a WMAP-7 best-fit model. Both simulations have a volume of  $(\sim 250 h^{-1} \text{ Mpc})^3$  and  $1024^3$  particles. Details of the simulations, including the values for the cosmology parameters, are provided in Section 2. The main difference between the two cosmologies is reflected in their values for  $\Omega_{\text{cdm}}$  and  $\sigma_8$ . Both parameters are larger for the Planck-18 cosmology, leading to enhanced structure formation. In addition, the value for  $h$  is smaller in the Planck-18 cosmology compared to WMAP-7. As for our previous tests, we investigate the core mass function for three different host halo mass bins. We apply the same mass model parameters to the two simulations.

Figure 8 summarizes our findings. We show the core mass functions themselves and the ratio for the two cosmologies in each host halo mass bin. The agreement is very good and shows that the approach is insensitive with regard to cosmology, at least in the family of  $\Lambda$ CDM models close to the current best-fit cosmology. Van den Bosch & Jiang (2016) find similarly that the effect of cosmology on the SHMF is negligible for models close to the current best-fit  $\Lambda$ CDM model.

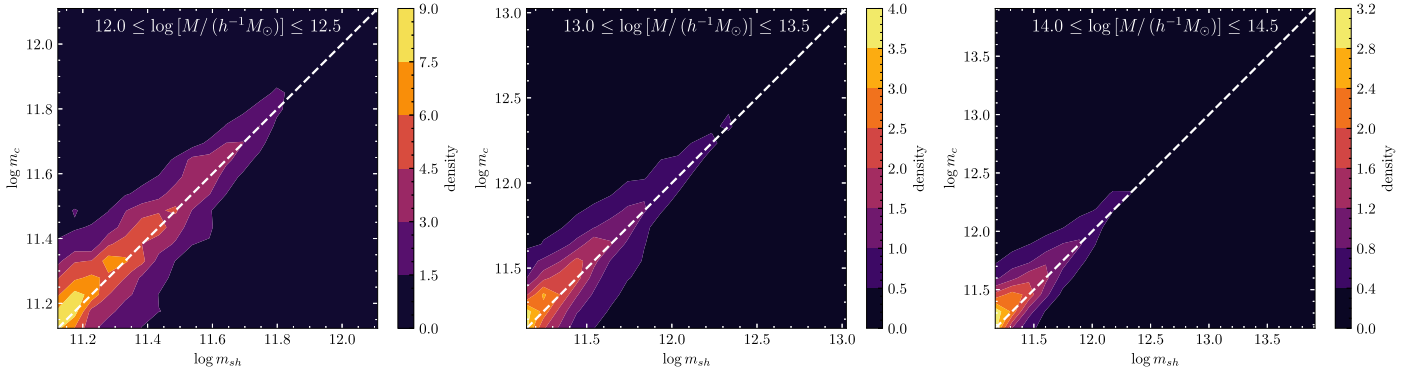
## 6. Last Journey Core Mass Function

Following the description and validation of the SMACC approach above, we now apply the model to the main Last Journey simulation. The simulation evolved  $10,752^3$  particles in a  $(3400 h^{-1} \text{ Mpc})^3$  volume, leading to a mass resolution of  $m_p \sim 2.721 \times 10^9 h^{-1} M_{\odot}$ . The volume covered is  $\sim 2500$  times the size of the SV simulation. A detailed description of the Last Journey simulation, including first-level analysis results such as matter power spectrum and mass function measurements, is provided in Heitmann et al. (2021).

The top panels of Figure 9 show the resulting core mass functions at redshifts  $z = 0$  (left panel) and  $z = 1$  (right panel) for five host halo mass bins. As a consequence of the excellent statistics available for the large Last Journey simulation, we have chosen very narrow mass bins that are much smaller in mass width than the mass bins used in the earlier sections for the smaller simulations. Table 3 lists the mass bins and the host



**Figure 6.** Distribution of the relative difference in mass of spatially matched subhalos and cores in SV. The subset of objects shown are subhalos matched to the most massive core within a search radius of  $2r_{\text{sh,vir}}$  at  $z = 0$ , where we only consider subhalos and cores with  $m > 100$  particles for the matching. For the three host mass bins shown, the fractions of matched subhalos (i.e., subhalos that have at least one core above the mass cut within the search radius) are 74% (left panel), 85% (middle panel), and 90% (right panel) of the total subhalos above the mass cut. The relative difference between subhalo mass and core-evolved mass (both at  $z = 0$ ) is shown in red, where the fiducial model parameters were used to evolve core mass. The relative difference between subhalo mass (at  $z = 0$ ) and core infall mass (i.e., the mass of the halo before merging) is shown in blue.



**Figure 7.** Mass comparison contours at  $z = 0$  of spatially matched subhalos and cores in SV with  $m > 100$  particles as in Figure 6. Subhalo mass  $m_{\text{sh}}$  is plotted against core-evolved mass  $m_c$  for three host halo mass bins. The dashed diagonal lines indicate  $m_c = m_{\text{sh}}$ .

halo counts for both redshifts. Our chosen parameter values of  $(\mathcal{A}, \zeta) = (1.1, 0.1)$  were used to evaluate the mass-loss model (see Section 4.2 for a discussion of the optimal parameters). As examined in Section 4.3, following the convergence results established with the HM and SV simulations, we impose a mass cut for the core mass for the Last Journey simulation at 100 particles. We consider cores above this threshold to reliably model substructure mass evolution, and they can serve as input for a future SAM project. To illustrate this mass cut, points for which  $\log(m/M) \leq \log(100m_{p,\text{LJ}}/\langle M \rangle)$  are shown with a higher color transparency ( $m_{p,\text{LJ}}$  is the particle mass of the Last Journey simulation;  $\langle M \rangle$  is the average host halo mass of each bin).

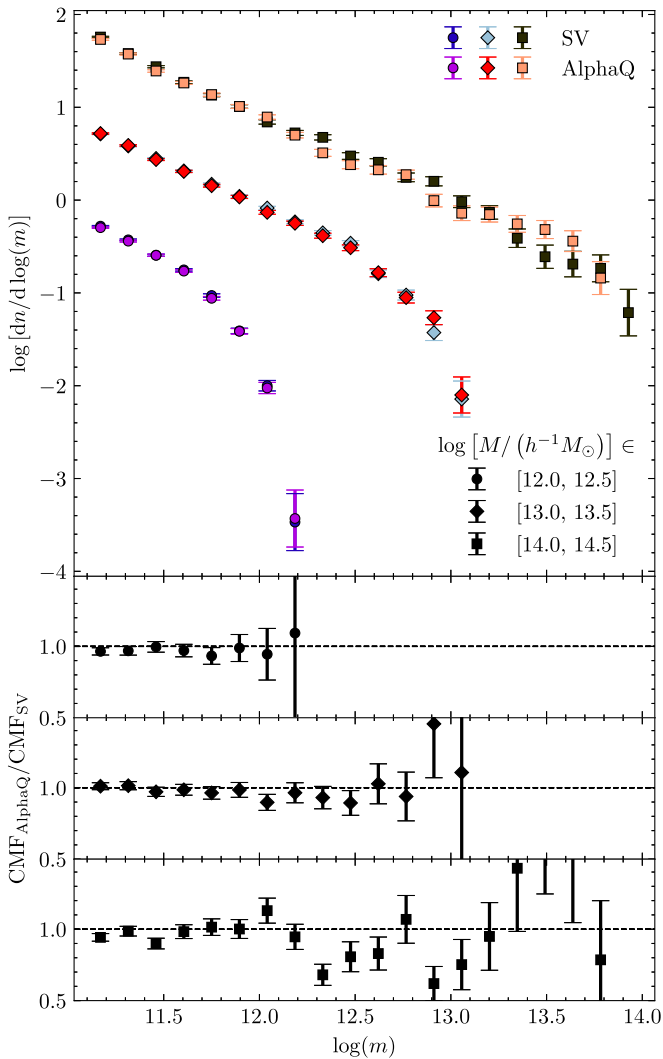
For each redshift, to enable comparison of our mass functions with other studies, we fit Equation (A1) to our core mass function in each host halo mass bin. In van den Bosch & Jiang (2016), the authors showed that SHMFs determined at  $z = 0$  using the Bolshoi and MultiDark simulations and the output of two different subhalo finders agree with this fitting form for  $\log(m/M) \lesssim -1$ .<sup>7</sup> In our results, we consider points for which  $\log(m/M) > \log(100m_{p,\text{LJ}}/\langle M \rangle)$  (i.e., the opaque points shown in Figure 9) when fitting to each mass bin.

We first fit Equation (A1) with both  $\alpha$  and  $A_M$  as free parameters, and we found the best-fit slope values at  $z = 0$  of  $\alpha = 0.849, 0.953, 0.929, 0.928$ , and  $0.939$  for the  $\log[M/(h^{-1}M_\odot)] = 12.5, 13.0, 13.5, 14.0$ , and  $14.5$  bins, respectively. At  $z = 1$ , the best-fit values are  $0.883, 0.852, 0.776, 0.770$ , and  $0.797$  for the same bins. We empirically selected  $\alpha = 0.94$  ( $z = 0$ ) and  $\alpha = 0.8$  ( $z = 1$ ) as the overall best-fit values for the slopes that provide reasonable fits for all of our host halo mass bins. Holding  $\alpha$  fixed for each redshift, we refit the amplitude  $A_M$  from Equation (A1) to our mass function for each bin. The bottom panels of Figure 9 show the ratio of our core mass function to the resulting fitting function for each redshift and host halo mass bin. The best-fit parameter value for the normalization  $A_M$  is displayed in each panel.

Equation (A1), with slope  $\alpha$  fixed and the normalization  $A_M$  allowed to vary, fits the data points well for both redshifts and for each mass bin for which  $\log(m/M) > \log(100m_{p,\text{LJ}}/\langle M \rangle)$ . This result increases our confidence in considering cores with a mass above a limit of 100 particles to be reliable models of substructure (see Section 4.3). Moreover, we confirmed that fitting the core mass function for both smaller simulations (SV and HM) yielded similar results to our presented values, further demonstrating consistency between the runs.

In the literature, slope values in the range of about 0.7–1.1 have been found by fitting the SHMF computed from  $N$ -body

<sup>7</sup> We note that van den Bosch & Jiang (2016) use the virial mass definition for host halo mass while we use the FOF,  $b = 0.168$  mass.



**Figure 8.** Comparison of the mass model applied to two simulations with different cosmologies. Top panel: average core mass functions at  $z=0$  for three host halo mass bins in the SV and AlphaQ simulations. Note that the  $x$ -axis is the core mass  $m$  and that we consider only cores for which  $m > 100m_{p,SV}$ , where  $m_{p,SV}$  is the particle mass of the SV simulation. The SV host bins contain 40,628, 4797, and 337 host halos, while the AlphaQ bins contain 37,182, 4335, and 286 halos. The host halo bins cover a mass range of  $[12.0, 12.5]$  (circles),  $[13.0, 13.5]$  (diamonds), and  $[14.0, 14.5]$  (squares) in  $\log[M/(h^{-1}M_{\odot})]$ . Bottom panels: ratio of the core mass functions shown in the top panel (AlphaQ to SV for each host halo mass bin).

simulations (Ghigna et al. 2000; Helmi et al. 2002; De Lucia et al. 2004; Shaw et al. 2006; Diemand et al. 2007; Angulo et al. 2009; Boylan-Kolchin et al. 2010; Giocoli et al. 2010; Gao et al. 2012). We note that a number of factors can possibly lead to variations in the slope values, including differences in the host halo mass definition (we use an FOF,  $b = 0.168$  mass, while some others have used  $M_{200}$  or virial mass definitions), differences in the form of the power-law fitting function, and differences in the definition of the SHMF (e.g., cumulative and differential SHMFs).

One important consideration when comparing our slope measurements comes from our mass-loss model parameter selection. As described in Section 4.2, we chose parameters that provide the best overall fit to the SHMF at redshifts  $z=0$  and  $z=1$ . This result provided reasonable results for both redshifts. However, there was a measured variation of best-fit

parameters when individually fitting each redshift independently (indicating a slight redshift dependence that was not fully captured by our simple model). We have repeated the slope measurements at  $z=0$  for the HM and SV simulations using mass model parameter values that provide an optimal fit for our  $z=0$  SHMFs. We found the fitted slopes of some mass bins to be closer to the value given in Jiang & van den Bosch (2016) of  $\alpha=0.86$ , as expected given our measured subhalo finder agreement in Appendix A. Nevertheless, the empirical measurements we found of  $\alpha=0.94$  ( $z=0$ ) and  $\alpha=0.8$  ( $z=1$ ) using our fiducial model parameters are consistent with the range in past works and allow us to maintain a simple self-contained model without the addition of redshift-dependent modifications.

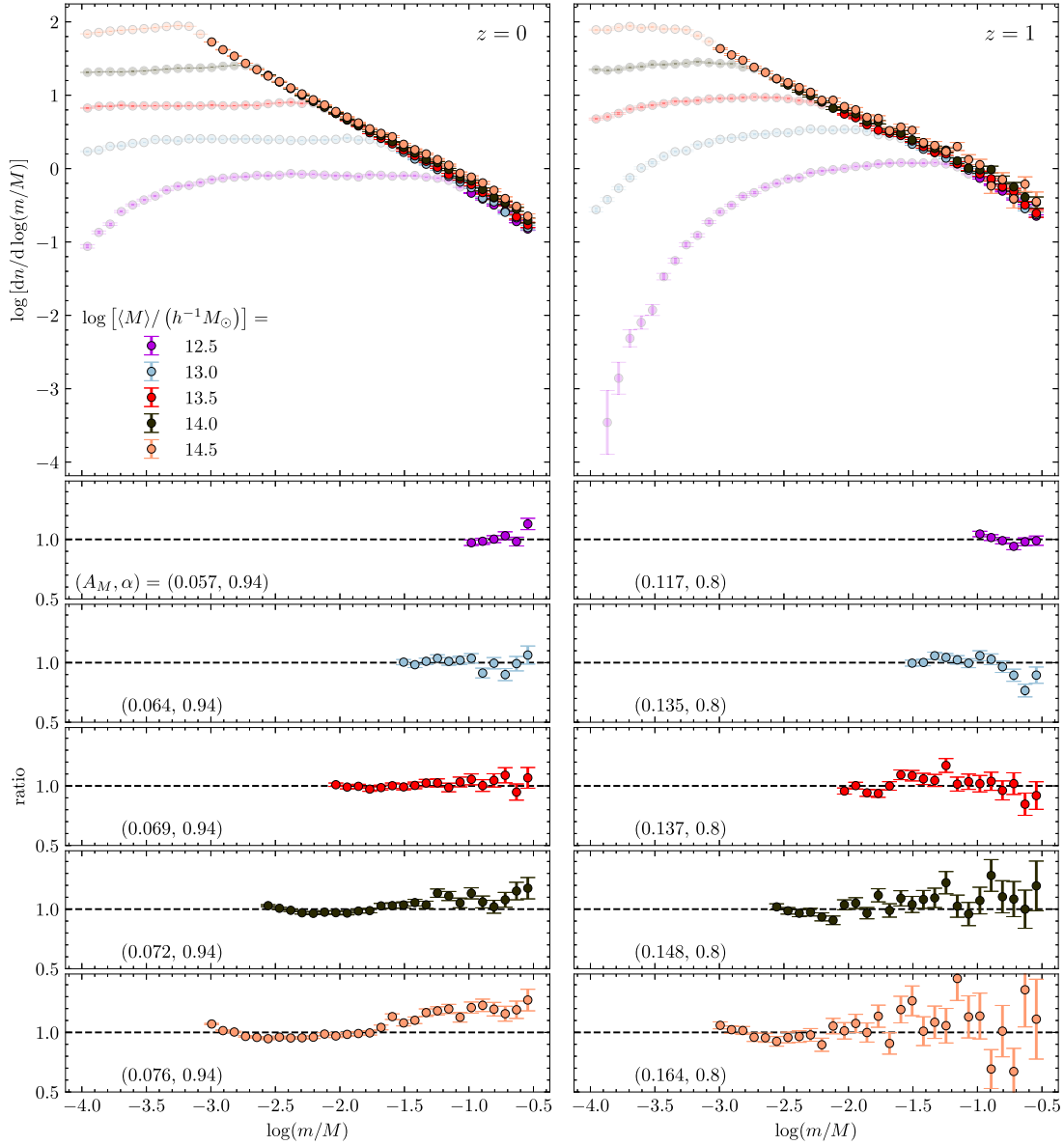
## 7. Summary and Outlook

In this paper, we introduced SMACC, a methodology for enabling the use of halo cores as an efficient substitute for subhalo catalog generation. SMACC provides mass estimates for halos after they have merged to become substructures within other halos. The mass-loss model uses the infall mass of a merging halo as a starting point and then evolves the mass via a simple two-parameter model. The core-tracking approach (summarized in Section 3) in combination with SMACC provides a very fast, scalable method to follow the evolution of substructure within halos over time. We apply the SMACC procedure to the Last Journey simulation (reported in Section 6), which is described in the first paper of the Last Journey series (Heitmann et al. 2021).

We established our approach in Section 4, where we fit results obtained using SMACC to the SHMF measured on the same set of halos, allowing us to choose best-fit parameters where the model agreed well with the subhalo results at  $z=0$  and  $z=1$ . (The results from our subhalo finder are consistent with those from other commonly used finders, as shown in Appendix A.) Importantly, we carried out mass resolution tests (Section 4.3) to demonstrate the validity of our approach with respect to mass thresholds, as well as a range of efficacy tests (Section 5). These tests included the investigation of the robustness of SMACC toward different host halo properties, in particular the host halo concentration as discussed in Section 5.1; we found that the simple mass-loss model holds for different host halo subpopulations based on concentration. We also carried out an individual subhalo-to-core matching comparison in Section 5.2, showing good agreement between the measurements. Finally, in Section 5.3 we tested the robustness of our approach against changes in the cosmological models as currently allowed by observational constraints and found essentially no cosmology dependence.

One of the primary aims for the development of SMACC is the generation of merger trees with embedded substructure information acquired through core tracking. These merger trees can then be used in many different science studies, such as input to SAMs or empirical models for the galaxy–halo connection that rely on using detailed information of the evolution of halo substructure over time. In this paper, we have focused on demonstrating that SMACC can successfully deliver an alternative approach for building detailed halo substructure models. We have taken the first step of subhalo mass modeling and have not considered the treatment of core mergers as potential proxies for galaxy mergers and of core disruption as a possible proxy for galaxy disruption. The





**Figure 9.** Results from the main Last Journey simulation for the core mass model. Top panels: average core mass functions at  $z=0$  (left) and  $z=1$  (right) for five host halo mass bins. (See Table 3 for descriptions of the host mass bins.) To illustrate the core mass cut of 100 particles (see Section 4.3), points for which  $\log(m/M) \leq \log(100m_{p,LJ}/\langle M \rangle)$  have enhanced color transparency, where  $m_{p,LJ}$  is the particle mass of the Last Journey simulation and  $\langle M \rangle$  is the average host halo mass of each host halo bin. For each host halo mass bin, we fit Equation (A1) to the opaque data points of that bin, first fixing  $\alpha = (0.94, 0.8)$  for  $z = (0, 1)$  and finding a best-fit value for  $A_M$ . Bottom panels: for each host halo mass bin, the ratio of the opaque section of the core mass function (shown in the top panel) to the fitting function, for both redshifts. The best-fit parameter value for  $A_M$  is displayed in each panel.

**Table 3**  
Last Journey Host Halo Mass Bins

$\log[\langle M \rangle / (h^{-1} M_\odot)]$	$M_{\min} (10^{12} h^{-1} M_\odot)$	$M_{\max} (10^{12} h^{-1} M_\odot)$	Host Halo Count
12.5 (12.5)	3.163 (3.163)	3.163 (3.163)	42,640 (32,902)
13.0 (13.0)	10.000 (10.000)	10.005 (10.005)	14,303 (8365)
13.5 (13.5)	31.625 (31.625)	31.679 (31.666)	10,102 (2929)
14.0 (14.0)	100.000 (100.000)	100.674 (100.438)	10,018 (1093)
14.5 (14.5)	316.229 (316.245)	328.172 (331.645)	10,001 (354)

**Note.** Halo masses and counts for the Last Journey simulation. The numbers given in each column correspond to the values for the  $z=0$  host halo bins and are followed by numbers in parentheses that correspond to the values for  $z=1$  host halo bins. The bins are used in Figure 9 for the mass-loss model applied to the large Last Journey simulation. As a consequence of the excellent statistics available for the large Last Journey simulation, we have chosen very narrow mass bins that have a much smaller width than the bins used in the earlier sections for the smaller simulations.

availability of information on the distribution of particles within a core can be applied to monitor such processes. In the specific context of modeling the galaxy distribution in clusters, D. Korytov et al. (2021, in preparation) describe a core merging and disruption procedure that utilizes core particle distribution data. We will incorporate and further develop such an approach in future work to provide more simulation-based inputs to semianalytic modeling approaches for galaxy formation.

We thank Andrew Hearin and Malin Renneby for many useful discussions and Claude-André Faucher-Giguère for valuable comments. We acknowledge important contributions from Hal Finkel, Adrian Pope, and Tom Uram for previous work on HACC and the first paper in the Last Journey series. Argonne National Laboratory’s work was supported under the U.S. Department of Energy contract DE-AC02-06CH11357. Awards of computer time were provided by the ASCR Leadership Computing Challenge (ALCC) program. This research used resources of the Argonne Leadership Computing Facility at the Argonne National Laboratory, which is supported by the Office of Science of the U.S. Department of Energy under contract No. DE-AC02-06CH11357. We are indebted to the ALCF team for their outstanding support and help in enabling us to carry out a simulation at this scale. This research also used resources of the Oak Ridge Leadership Computing Facility, which is a DOE Office of Science User Facility supported under contract DE-AC05-00OR22725.

## Appendix A

### Subhalo Finder Description and Validation

In order to develop our mass modeling approach on cores and optimize the associated modeling parameters, we use results obtained from an independent subhalo finder run on the same simulation. We make use of the subhalo finder operating within the HACC analysis framework, CosmoTools, as described in Li et al. (2016), where it was used to identify substructures for detailed simulation-based strong-lensing studies. This subhalo finding approach is based on local density estimation (similar to, e.g., SUBFIND by Springel et al. 2001) but is then augmented by velocity information, helping to better distinguish whether a particle belongs to a subhalo or is a member of the host halo.

We first identify all halos in the simulation with a fast, parallel FOF finder with a linking length of  $b = 0.168$  (smaller than the often-used fiducial value of  $b = 0.2$  to reduce overlinking artifacts). The FOF halo finding approach, based on an approximate isodensity boundary, has a long history in cosmology and was first introduced by Davis et al. (1985). It is algorithmically well defined and useful down to small particle counts.

Following the halo finding step, we organize the particles that belong to the FOF host halo into a Barnes–Hut tree structure (Barnes & Hut 1986) to enable efficient access. The local mass density is then estimated for each particle using a smoothed particle hydrodynamics (SPH) kernel. Particles are assigned to subhalo candidates, and we build two initial membership lists, one for the main halo and one for the subhalo candidates.

For each candidate subhalo we determine whether all assigned subhalo particles are bound to the substructure. For this step we calculate the escape velocity for each particle with

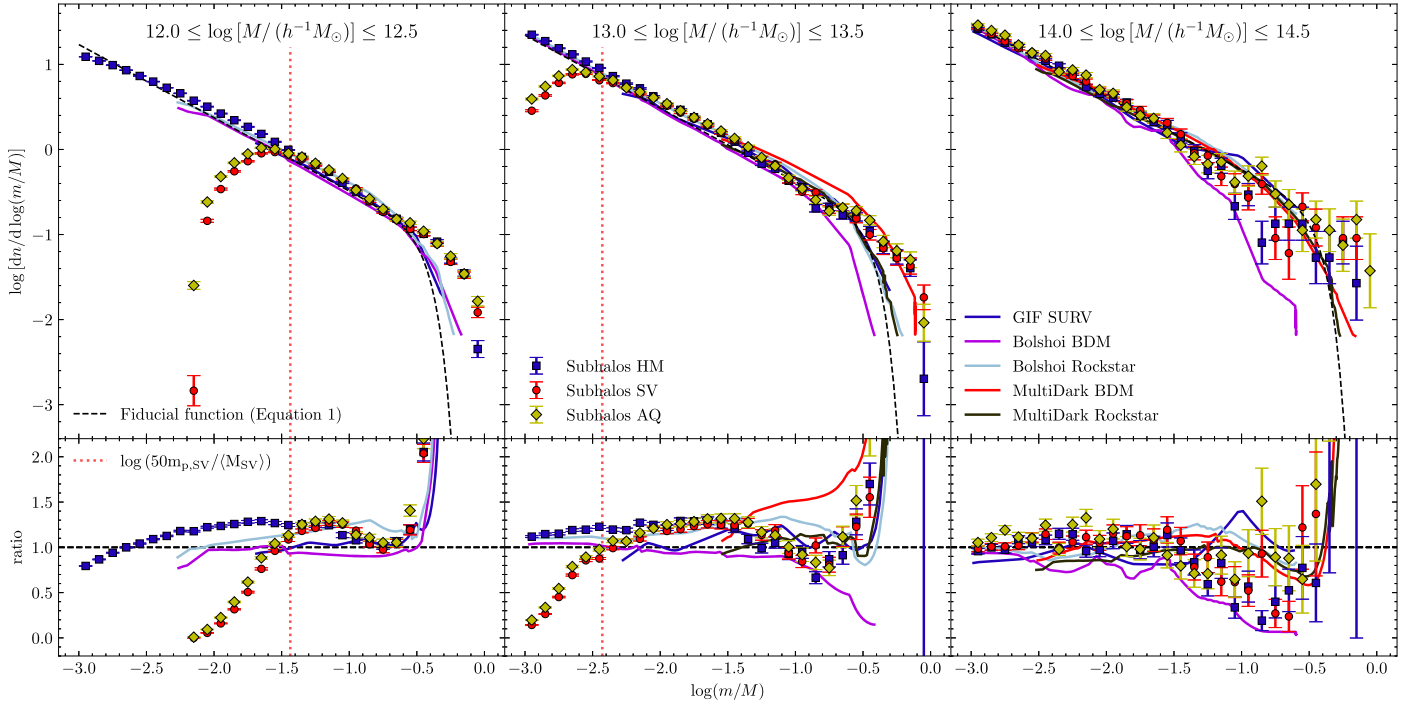
regard to the subhalo, using its total energy. If the particle velocity exceeds the escape velocity, we assign it to the main halo. A new energy calculation is then carried out and the process repeated until only bound particles remain in each subhalo. Finally, we use a mass threshold to determine whether the resulting subhalo candidate is still viable. Subhalos that fall below the preset minimum threshold for particle number are discarded. For each subhalo we record its properties in a subhalo property file (positions, velocities, velocity dispersions, and masses, as well as subhalo tags and tags to identify its host halo). We also separately save the particles belonging to the subhalos, including the relevant tag identifiers.

For comparison with other subhalo finders, we use the Last Journey-SV and HM simulations and the AlphaQ simulation at  $z = 0$ . We measure the SHMF for all three simulations in three mass bins and compare the results to previous findings. In particular, we take advantage of the useful compilation of SHMF measurements presented in van den Bosch & Jiang (2016). For the host halos, we impose a mass cut of just under  $10^{11} h^{-1} M_{\odot}$  (see Table 1 for the exact limits). The mass cut is chosen to avoid mass resolution effects in the subhalo finder for the Last Journey-SV and AlphaQ simulations. For the comparison, we also measure the  $M_{200c}$  host halo mass and convert it to a virial mass estimate using halo concentration measurements. We identify subhalos with over 20 particles. This last requirement leads to a minimum subhalo mass of over  $\sim 2.5 \times 10^{10} h^{-1} M_{\odot}$  and  $\sim 9.3 \times 10^8 h^{-1} M_{\odot}$  for the SV and HM simulations, respectively, and  $\sim 2.4 \times 10^{10} h^{-1} M_{\odot}$  for the AlphaQ simulation. The HM simulation allows us to investigate the convergence of the SHMF down to very low masses. We will use the subhalo finder results for  $z = 1$  and  $z = 0$  for the SV and HM simulations (with FOF host mass measurements using  $b = 0.168$ ) throughout the paper to tune and verify our mass modeling approach for cores.

The top panels of Figure 10 compare our measurements with SHMFs delivered by different subhalo finders in three mass bins, while the bottom panels show the ratios with respect to the fiducial SHMF given in Equation (6) of van den Bosch & Jiang (2016):

$$\frac{dN}{d \log(m/M)} = A_M \left( \frac{m}{M} \right)^{-\alpha} \exp[-50(m/M)^4], \quad (\text{A1})$$

with  $A_M = 0.045, 0.058$ , and  $0.080$  for the  $\log[M/(h^{-1} M_{\odot})] \in [12.0, 12.5], [13.0, 13.5]$ , and  $[14.0, 14.5]$  host halo bins, respectively, and  $\alpha = 0.86$  as in their Figure 5. We indicate the fiducial functions in the top panels with a dashed line. The symbols with error bars are from our simulations (for the four largest FOF halos from the HM simulation shown in the right panel in the figure, we use an approximate unbinding algorithm to speed up the calculation). The left, middle, and right panels of the figure show bins containing (42,183, 41,128, 36,524), (4965, 4929, 4354), and (373, 331, 267) host halos for the (HM, SV, AlphaQ) simulations, respectively. The solid lines in Figure 10 are from van den Bosch & Jiang (2016). We note the generally good agreement of our SHMF with those measured with commonly used subhalo finders. The results are consistent with those reported in the major subhalo finder comparison carried out in Onions et al. (2012), which found overall agreement for the SHMF at the 20% level; van den Bosch & Jiang (2016) confirmed these findings as well.



**Figure 10.** Comparison of our subhalo finder results to those from other implementations. Top panels: average SHMF at  $z = 0$  for three host halo mass bins in the HM (squares), SV (circles), and AlphaQ simulations (diamonds) where subhalos with  $m > 20$  particles are shown (see text for the number of host halos in each bin). Results from a number of other simulations using different subhalo finders (solid lines) are shown as reported in van den Bosch & Jiang (2016). Following van den Bosch & Jiang (2016), we use Equation (A1) with  $\alpha = 0.86$  and  $A_M = 0.045, 0.058$ , and  $0.080$  for the  $\log [M/(h^{-1}M_\odot)] \in [12.0, 12.5], [13.0, 13.5]$ , and  $[14.0, 14.5]$  host bins, respectively, as fiducial fitting functions to compare the results of each bin. The fiducial fitting functions are indicated with dashed lines. The vertical dotted line in the left and middle panels approximates the 50-particle mass threshold for SV, following van den Bosch & Jiang (2016). Bottom panels: ratio of the SHMFs to the fiducial fitting functions (dashed lines) shown in the top panels. Note that while we use the FOF finder to identify the host halos, for this plot we have used the corresponding virial mass measurement for determining  $M$  to enable a meaningful comparison with the other results.

There are two small differences between the measurement carried out in van den Bosch & Jiang (2016) and our analysis. First, van den Bosch & Jiang (2016) use the virial mass for the host halo mass measurement, which is close to  $M_{100c}$  and leads to higher host halo estimates compared to the FOF,  $b = 0.168$  host halo mass used here. In order to compensate for this difference, we use a virial mass estimate for the host halo mass when presenting our results. We note that since  $M_{\text{vir}}$  halos are larger than FOF,  $b = 0.168$  halos, subhalos on the outskirts might be undercounted. Second, the cosmology used in our paper is different. The Bolshoi and MultiDark simulations are based on WMAP results. The Planck-18 parameters used in our paper have slightly higher values for  $\sigma_8$  and  $\Omega_{\text{cdm}}$  and lead to more clustering. However, the comparison of the SHMF from the SV and the AlphaQ simulations (Planck-18 and WMAP-7 cosmologies, respectively) in Figure 10 shows that varying  $\Lambda$ CDM parameters, while staying close to the current best-fit cosmology, has a negligible effect.

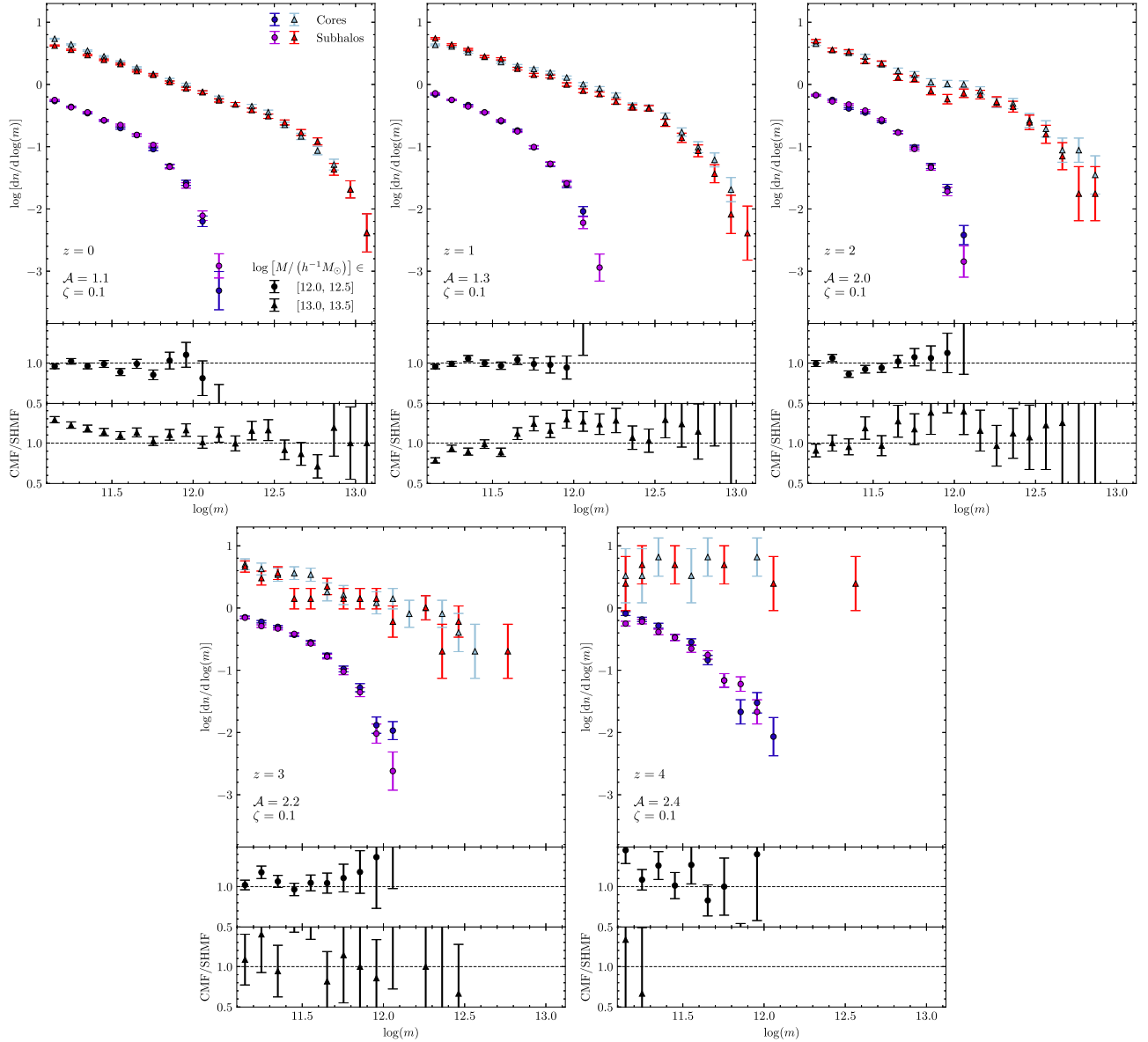
Besides establishing the overall agreement with other subhalo finders, the comparison between the SV and HM results is important for convergence studies. The excellent agreement of the SHMF in the (higher subhalo mass) overlap region shows that the lower mass resolution faithfully captures the subhalos of interest. In the lower range of subhalo masses, the limited mass resolution of the SV simulation results in subhalo loss and the SHMF rolls off. (We note that this resolution limit is well below the subhalo mass cutoff that we are planning to use for further studies with the main Last Journey simulation.)

## Appendix B Redshift-dependent Mass Modeling

Throughout the paper we have focused on results for redshifts  $z = 0$  and  $z = 1$ . This allows us to make comparisons to the existing literature for, e.g., the subhalo finder. In addition, most observational results from cosmological surveys, such as SDSSs, BOSS, eBOSS, DES, and in the near future DESI, are provided in that redshift range, making it the most relevant for our purposes. Another, more technical reason for this focus is the low halo number statistics for our test simulations at high redshifts. Not only does the number of possible host halos rapidly decline with redshift, but the number of viable subhalos also decreases quickly.

Nevertheless, it is important to investigate the validity of the mass modeling approach at higher redshifts. As we note in Section 4.2, for  $z = 0$  and  $z = 1$  we already observe a small redshift dependence for the mass modeling best-fit parameters. In order to investigate this further, we ran the subhalo finder on the SV simulation for three additional redshifts,  $z = 2, 3$ , and  $4$ . We applied the mass model with the fiducial ( $\mathcal{A}, \zeta$ ) values, chosen to be  $(1.1, 0.1)$ , to the core catalog and compared the resulting SHMF to the core mass function. We found that the shape of the mass function was well captured with our choice for  $\zeta$  but the amplitude match got worse with higher redshift—the core mass function results were too high compared to the SHMF.

We therefore refit  $\mathcal{A}$  for all five available redshifts, keeping  $\zeta = 0.1$ . We focus our investigation on two host halo mass bins,  $\log [M/(h^{-1}M_\odot)] \in [12.0, 12.5]$  and  $[13.0, 13.5]$ . The halo



**Figure 11.** Comparison of the SHMFs to the core mass functions at  $z=0$  to  $z=4$ . At each redshift, we fit for the mass model parameter  $\mathcal{A}$  in the  $\log[M/(h^{-1}M_{\odot})] \in [12.0, 12.5]$  mass bin while keeping  $\zeta$  fixed at the chosen value (see Section 4.2). The results for each redshift show the best-fit  $\mathcal{A}$  value that was found; the value is indicated in each top panel. Top panels: average subhalo and core mass functions for two host halo mass bins in the SV simulation. Subhalos and cores with  $m > 100$  particles are shown. Bottom panels: ratio of the core mass function to the SHMF for the two host halo mass bins.




count in the cluster mass bin  $[14.0, 14.5]$  that we consider in the main paper is very low at high redshift, and the results are therefore very noisy. Figure 11 shows the final results for the subhalo and core mass functions using the newly fitted value for  $\mathcal{A}$  for each redshift. To determine  $\mathcal{A}$ , we carried out a fitting procedure as described in Section 4.2 (using  $N_p = 1$  since  $\zeta$  was fixed at the fiducial value) for the  $\log[M/(h^{-1}M_{\odot})] \in [12.0, 12.5]$  mass bin. (The best-fit values for the higher-mass bin were found to be very similar.) We note that the best-fit value rapidly increases with redshift. A linear fit for  $\mathcal{A}$  as a function of redshift can be found:  $\mathcal{A}(z) = 0.35z + 1.1$ . The fit provides a perfect match for  $z=0$  and values above and below the best-fit values for all other redshifts. In order to derive a more robust fit, larger simulation volumes would be required that would provide

better statistics. We note that Hiroshima et al. (2018) provide redshift-dependent fits for both amplitude and slope of their mass model. A direct comparison to our result is difficult, however, since the functional form of their mass model differs from ours.

To summarize, we investigated the redshift dependence of the optimal mass model that provides a good fit between the subhalo and core mass function. We study the mass functions out to redshift  $z=4$ . The best-fit value governing the shape of the mass model changes minimally, while the amplitude increases at higher redshift. The derivation of a  $z$ -dependent mass-loss model is possible but beyond the scope of this paper. For the redshift range between  $z=0$  and  $z=1$  considered throughout the paper, a fixed value for both slope and amplitude provides a good description of the results.



## ORCID iDs

Imran Sultan  <https://orcid.org/0000-0003-2341-1534>  
 Katrin Heitmann  <https://orcid.org/0000-0003-1468-8232>  
 Eve Kovacs  <https://orcid.org/0000-0002-2545-1989>

## References

- Angulo, R. E., Lacey, C. G., Baugh, C. M., & Frenk, C. S. 2009, *MNRAS*, **399**, 983
- Barnes, J., & Hut, P. 1986, *Natur*, **324**, 446
- Behroozi, P. S., Wechsler, R. H., Wu, H.-Y., Busha, M. T., Klypin, A. A., & Primack, J. R. 2013, *ApJ*, **763**, 18
- Bertschinger, E. 1998, *ARA&A*, **36**, 599
- Boylan-Kolchin, M., Springel, V., White, S. D. M., & Jenkins, A. 2010, *MNRAS*, **406**, 896
- Bryan, G. L., & Norman, M. L. 1998, *ApJ*, **495**, 80
- Budzynski, J. M., Kopeck, S. E., McCarthy, I. G., McGee, S. L., & Belokurov, V. 2012, *MNRAS*, **423**, 104
- Campbell, D., van den Bosch, F. C., Padmanabhan, N., et al. 2018, *MNRAS*, **477**, 359
- Child, H. L., Habib, S., Heitmann, K., et al. 2018, *ApJ*, **859**, 55
- Davis, M., Efstathiou, G., Frenk, C. S., & White, S. D. M. 1985, *ApJ*, **292**, 371
- De Lucia, G., Kauffmann, G., Springel, V., et al. 2004, *MNRAS*, **348**, 333
- Diemand, J., Kuhlen, M., & Madau, P. 2007, *ApJ*, **667**, 859
- Dolag, K., Borgani, S., Schindler, S., et al. 2008, *SSRv*, **134**, 229
- Fakhouri, O., & Ma, C.-P. 2008, *MNRAS*, **386**, 577
- Frenk, C. S., & White, S. D. M. 2012, *AnP*, **524**, 507
- Gao, L., Navarro, J. F., Frenk, C. S., et al. 2012, *MNRAS*, **425**, 2169
- Gao, L., White, S. D. M., Jenkins, A., Stoehr, F., & Springel, V. 2004, *MNRAS*, **355**, 819
- Ghigna, S., Moore, B., Governato, F., et al. 2000, *ApJ*, **544**, 616
- Giocoli, C., Tormen, G., Sheth, R. K., & van den Bosch, F. C. 2010, *MNRAS*, **404**, 502
- Giocoli, C., Tormen, G., & van den Bosch, F. C. 2008, *MNRAS*, **386**, 2135
- Guo, Q., & White, S. 2014, *MNRAS*, **437**, 3228
- Habib, S., Pope, A., Finkel, H., Frontiere, N., et al. 2016, *NewA*, **42**, 49
- Han, J., Cole, S., Frenk, C. S., Benitez-Llambay, A., & Helly, J. 2018, *MNRAS*, **474**, 604
- Heitmann, K., Frontiere, N., Rangel, E., et al. 2021, *ApJS*, **252**, 19
- Helmi, A., White, S. D. M., & Springel, V. 2002, *PhRvD*, **66**, 063502
- Hiroshima, N., Shin'ichiro, A., & Ishiyama, T. 2018, *PhRvD*, **97**, 123002
- Jiang, F., & van den Bosch, F. C. 2014, *MNRAS*, **440**, 193
- Jiang, F., & van den Bosch, F. C. 2016, *MNRAS*, **458**, 2848
- Kaiser, N. 1984, *ApJ*, **284L**, 9
- Komatsu, E., Smith, K. M., Dunkley, J., et al. 2011, *ApJS*, **192**, 18
- Korytov, D., Hearin, A., Kovacs, E., et al. 2019, *ApJS*, **245**, 26
- Lacey, C. G., & Cole, S. 1993, *MNRAS*, **262**, 627
- Li, N., Gladders, M. D., Rangel, E. M., et al. 2016, *ApJ*, **828**, 54
- Lin, Y. T., Mohr, J. J., & Stanford, S. A. 2004, *ApJ*, **610**, 745
- Mo, H., van den Bosch, F., & White, S. D. M. 2011, *Galaxy Formation and Evolution* (Cambridge: Cambridge Univ. Press)
- Muldrew, S. I., Pearce, F. R., & Power, C. 2011, *MNRAS*, **410**, 2617
- Navarro, J. F., Frenk, C. S., & White, S. D. M. 1996, *ApJ*, **462**, 563
- Navarro, J. F., Frenk, C. S., & White, S. D. M. 1997, *ApJ*, **490**, 493
- Onions, J., Knebe, A., Pearce, F. R., et al. 2012, *MNRAS*, **423**, 1200
- Planck Collaboration, Aghanim, N., & Akrami, Y. 2020, *A&A*, **641**, A6
- Rangel, E., Li, N., Habib, S., et al. 2016, in *IEEE Int. Conf. on Cluster Computing (CLUSTERS)* (Piscataway, NJ: IEEE), 30
- Rodriguez-Gomez, V., Genel, S., Vogelsberger, M., et al. 2015, *MNRAS*, **449**, 49
- Shaw, L. D., Weller, J., Ostriker, J. P., & Bode, P. 2006, *ApJ*, **646**, 815
- Somerville, R. S., & Davé, R. 2015, *ARA&A*, **53**, 51
- Springel, V., White, S. D. M., Tormen, G., & Kauffmann, G. 2001, *MNRAS*, **328**, 726
- van den Bosch, F. C. 2017, *MNRAS*, **468**, 885
- van den Bosch, F. C., & Jiang, F. 2016, *MNRAS*, **458**, 2870
- van den Bosch, F. C., & Ogiya, G. 2018, *MNRAS*, **475**, 4066
- van den Bosch, F. C., Tormen, G., & Giocoli, C. 2005, *MNRAS*, **359**, 1029
- Wechsler, R., & Tinker, J. 2018, *ARA&A*, **56**, 435
- Wechsler, R. H., Bullock, J. S., Primack, J. R., Kravtsov, A. V., & Dekel, A. 2002, *ApJ*, **568**, 52
- Wechsler, R. H., Zentner, A. R., Bullock, J. S., & Kravtsov, A. V. 2006, *ApJ*, **652**, 71
- White, S. D. M., Davis, M., Efstathiou, G., & Frenk, C. S. 1987, *Natur*, **330**, 451
- Zentner, A. R. 2007, *IJMPD*, **D16**, 763
- Zentner, A. R., Berlind, A. A., Bullock, J. S., Kravtsov, A. V., & Wechsler, R. H. 2005, *ApJ*, **624**, 505